

# Matrix-valued Bessel processes

Martin Larsson

Swiss Finance Institute, Ecole Polytechnique Fédérale de Lausanne

January 16, 2013

## Abstract

This paper introduces a matrix analog of the Bessel processes, taking values in the set  $E$  of real square matrices with nonnegative determinant. They are related to the well-known Wishart processes in a simple way: the latter are obtained from the former via the map  $x \mapsto x^\top x$ . The main focus is on existence and uniqueness, questions which lead us to develop several new results of potential theoretic nature concerning the space of real square matrices. Specifically, the function  $w(x) = |\det x|^\alpha$  is a weight function in the Muckenhoupt  $A_p$  class for  $-1 < \alpha \leq 0$  ( $p = 1$ ) and  $-1 < \alpha < p - 1$  ( $p > 1$ ). The set of matrices of co-rank at least two has zero capacity with respect to the measure  $m(dx) = |\det x|^\alpha dx$  if  $\alpha > -1$ , and if  $\alpha \geq 1$  this even holds for the set of all singular matrices. As a consequence we obtain density results for Sobolev spaces over  $E^o$  with Neumann boundary conditions. The proofs of these results rely crucially on the stratification of the state space into smooth manifolds consisting of fixed-rank matrices.

**Key words:** Matrix-valued process, Bessel process, Wishart process, Muckenhoupt weight, positive determinant matrix, reflecting boundary condition.

**Acknowledgment.** The author would like to thank Damir Filipović, Dmitry Drusvyatsky, and Jim Renegar for very stimulating discussions.

## 1 Introduction

The Wishart processes constitute a class of matrix-valued generalizations of the squared Bessel (BESQ) processes, and were studied systemically for the first time by Bru [4]. Her work has subsequently been extended in various directions, for instance by Donati-Martin et al. [7] and Cuchiero et al. [6], among others. The fact that the BESQ processes have well-behaved matrix-valued analogs begs the question of whether the same is true for the Bessel (BES) processes. Since Wishart processes take values in  $\mathbf{S}_+^d$ , the cone of positive semidefinite matrices, a natural attempt would be to consider their positive semidefinite square roots. However, as was already remarked by Bru [4], these processes appear intractable. Our goal is to show that by passing to a larger state space, a more

well-behaved class of processes can be obtained. Specifically, the state space will be the set

$$E = \{x \in \mathbf{M}^d : \det x \geq 0\},$$

where  $\mathbf{M}^d$  is the Euclidean space of all  $d \times d$  real matrices, endowed with the usual inner product  $x \bullet y = \text{tr}(x^\top y)$  and norm  $\|x\| = \sqrt{x \bullet x}$ . The *matrix-valued Bessel process with parameter  $\delta > 0$  and matrix dimension  $d$* , abbreviated BESM( $\delta, d$ ), will be an  $E$ -valued Markov process whose generator (acting on a suitable function space) is given by

$$\mathcal{A} = \frac{1}{2}\Delta + \frac{\delta - 1}{2}x^{-\top} \bullet \nabla, \quad (1)$$

where  $x^{-\top} = (x^{-1})^\top$ ,  $\nabla$  is the  $d \times d$  matrix with elements  $\partial_{x_{ij}}$  (so that for  $f \in C^1(\mathbf{M}^d)$ ,  $\nabla f$  is its gradient), and  $\Delta = \nabla \bullet \nabla = \sum_{i,j} \partial_{x_{ij}x_{ij}}^2$  is the Laplacian. Notice that  $\mathcal{A}$  is the generator of the BES( $\delta$ ) process when  $d = 1$ . A general discussion on BES and BESQ processes is available in [19, Chapter XI], and a specialized treatment of the BES( $\delta$ ) process,  $0 < \delta < 1$ , can be found in [2].

The singular behavior of the drift part of  $\mathcal{A}$  makes a treatment via the associated martingale problem inconvenient. Instead we use the machinery of Dirichlet forms, which turns out to be very suitable for proving existence. We will apply the theory as presented in the book by Fukushima et al. [12]. The crucial fact here is that  $\mathcal{A}$  is a symmetric operator with respect to the measure  $m(dx) = |\det x|^{\delta-1}dx$  (where  $dx$  is Lebesgue measure on  $\mathbf{M}^d$ ), which is a consequence of an integration by parts formula (Theorem 2). The precise definition of the BESM process and semigroup, together with the discussion of existence, is given in Section 2.

Uniqueness, discussed in Section 4, is a much more delicate issue, whose resolution occupies the majority of the paper. We will obtain Markov uniqueness of the associated semigroup, relying on density results for Sobolev spaces with Neumann boundary condition (Theorem 8). In the process we are lead to prove several results, interesting in their own right, about the measure  $m$  and its interaction with the state space. Specifically, we prove that  $|\det x|^\alpha$  is locally Lebesgue integrable on  $\mathbf{M}^d$  precisely when  $\alpha > -1$  (Theorem 1), and that it is a weight function in the Muckenhoupt  $A_p$  class when  $\alpha \in (-1, 0]$  and  $p = 1$ , and when  $\alpha \in (-1, p - 1)$  and  $p > 1$  (Theorem 9). This exactly parallels the situation for the weight function on  $\mathbf{R}$  given by  $t^\alpha$ . Moreover, we show that the matrices of co-rank at least two form a set of zero capacity with respect to  $m$ , and that the set of all singular matrices has zero capacity if  $\delta \geq 2$  (Theorem 10). The latter result relies on the fact that the sets

$$M_k = \{x \in \mathbf{M}^d : \text{rank } x = k\}$$

are smooth manifolds forming a stratification of  $\mathbf{M}^d$ . In particular, it allows us to deal with the highly non-convex, non-Lipschitz structure of the state space  $E$ . In fact, the interior  $E^\circ$  does not even lie on one side of its boundary  $\partial E$ , as can be seen by considering the lines  $\{tx : t \in \mathbf{R}\}$ ,  $x \in \mathbf{M}^d$  (which intersect  $\partial E$ ): If  $d$  is even, each line lies either entirely inside  $E$ , or entirely outside  $E^\circ$ . In the future we expect similar techniques to be useful for analyzing a wide class of processes on semialgebraic (or, more generally, stratifiable) state spaces. Some of the groundwork for this is laid in [8].

We emphasize that the case  $0 < \delta < 2$  is the difficult one, due to the fact that the boundary of the state space is reached almost surely. This is analogous to the scalar Bessel processes. When  $\delta \geq 2$ , on the other hand, the boundary of  $E$  is never reached, and the BESM process can be realized pathwise as the strong solution to a suitable stochastic differential equation. This is discussed at the end of Section 3.

Let us say something about why the BESM processes are natural analogs of the BES processes, other than the resemblance of their infinitesimal generators and the fact that they coincide for  $d = 1$ . The main reason is that the process  $X^\top X$ , where  $X$  is  $\text{BESM}(\delta, d)$ , is a Wishart process with parameter  $\alpha = d - 1 + \delta$ , denoted  $\text{WIS}(\alpha, d)$ . As a consequence,  $\|X\|$  is  $\text{BES}(d\alpha)$ , and  $\det X$  is a time-changed  $\text{BES}(\delta)$  process. These results are presented in Section 3. It is interesting to note that the restriction  $\alpha > d - 1$ , which corresponds exactly to the range where the symmetrizing measure  $m$  has the Radon property, is where the stochastic differential equation for the Wishart process has a weak solution, see [4, Theorem 2].

A second motivation, which was the original “clue” that led us to consider the generator  $\mathcal{A}$ , is as follows. Let  $X$  be an  $\mathbf{M}^d$ -valued Brownian motion starting from  $I$  (the identity matrix), and let  $\tau_0$  be the first time  $\det X_t$  hits zero. Then  $\det X_{t \wedge \tau_0}$  is a martingale, and we may use it to change the probability measure. An application of Girsanov’s theorem, using the identity  $\nabla \ln \det(x) = x^{-\top}$ , shows that under the new measure, the process

$$W_t = X_t - X_0 - \int_0^t X_s^{-\top} ds, \quad t \geq 0,$$

is Brownian motion. Thus  $X$  becomes a Markov process whose generator is  $\mathcal{A}$  with  $\delta = 3$ , and its determinant is positive by construction. This is fully analogous to the well-known construction of the  $\text{BES}(3)$  process as Brownian motion “conditioned to stay positive”.

In addition to the notation already introduced above, the following conventions will be in force throughout the paper.

- For any subset  $U \subset \mathbf{M}^d$ ,  $C(U)$  denotes the space of real-valued continuous function on  $U$ , equipped with the relative topology. The subspace of compactly supported functions is denoted by  $C_c(U)$ . If the functions take values in a space  $V$ , we write  $C(U; V)$ . A function on  $U$  is called *k times continuously differentiable* ( $k \in \mathbb{N} \cup \{\infty\}$ ) if around each  $x \in U$  it coincides with some  $k$  times continuously differentiable function on  $\mathbf{M}^d$ . The space of such functions is denoted  $C^k(U)$ . The meaning of  $C_c^k(U)$ ,  $C_c^k(U; V)$ , etc. should be clear. The following remark is trivial but important: the notion of “compactly supported” and “vanishing at infinity” depends on the class of compact sets, which in turn depends on  $U$ . In particular, if  $U$  is open in  $\mathbf{M}^d$ , a set  $K \subset U$  is compact in  $U$  if and only if it is bounded, closed in  $\mathbf{M}^d$ , and bounded away from  $\partial U$ . If on the other hand  $U$  is closed in  $\mathbf{M}^d$  (as is the case when  $U = E$ ), compact subsets *need not be bounded away from*  $\partial U$ . The same remark applies to any property that holds *locally on*  $U$ .
- As already mentioned, the sets  $M_k$ ,  $k = 1, \dots, d$ , constitute a stratification of  $\mathbf{M}^d$  into smooth manifolds, and the dimension of  $M_k$  is  $d^2 - (d - k)^2$ . Note that  $\partial E =$

$\cup_{k \leq d-1} M_k$ . The *tangent space* of  $M_k$  at  $x \in M_k$  is well-defined, and for any  $v \in \mathbf{M}^d$  we say that  $v$  is *tangent to  $\partial E$  at  $x \in \partial E$*  if  $v$  lies in the tangent space at  $x$  of the manifold  $M_k$  that contains  $x$ . We refer to [18] for background on differential geometry.

- $O(d)$  is the orthogonal group over  $\mathbf{R}^d$ , and  $T(d)$  is the group of upper-triangular  $d \times d$  real matrices with strictly positive diagonal entries. The set of nonsingular matrices (i.e., the general linear group) is homeomorphic to  $O(d) \times T(d)$  via the QR-decomposition.
- For  $x \in \mathbf{M}^d$ , we let  $\text{adj } x$  denote the *adjugate matrix* of  $x$  (i.e., the transpose of the matrix of cofactors). We will freely use the identities  $x \text{adj } x = (\det x)I$  and  $\nabla \det(x) = \text{adj } x$ , as well as the fact that  $\text{adj } x = 0$  if and only if  $\text{rank } x \leq d - 2$ .

## 2 Definition and existence

The definition of the BESM process departs from the differential operator  $\mathcal{A}$  in (1) acting on functions in  $C_{c,\perp}^2$ , where

$$C_{c,\perp}^2 = \left\{ f \in C_c^2(E) : x^{-\top} \bullet \nabla f \text{ is bounded} \right\}.$$

As we will see momentarily,  $(\mathcal{A}, C_{c,\perp}^2)$  is a symmetric operator on  $L^2(E, m)$ , where the measure  $m$  is given by

$$m(dx) = |\det x|^{\delta-1} dx.$$

Occasionally  $m$  will be viewed as a measure on  $\mathbf{M}^d$ , or on subsets other than  $E$ . It is crucial that  $m$  is a Radon measure on  $\mathbf{M}^d$ , and hence on  $E$ , when  $\delta > 0$ . The proof of this fact relies on the following lemma, which is an elegant consequence of the group structure associated with the QR-decomposition of an invertible matrix.

**Lemma 1** *Let  $f : \mathbf{M}^d \rightarrow \mathbf{R}$  be nonnegative and measurable. Then*

$$\int_{\mathbf{M}^d} f(x) |\det x|^{-d} dx = \int_{O(d) \times T(d)} f(QR) \mu(dQ) \prod_{i=1}^d R_{ii}^{-i} dR,$$

where  $\mu$  is a non-normalized Haar measure on  $O(d)$ , and  $dR = \prod_{i \leq j} dR_{ij}$ .

**Proof.** If  $f(x) |\det x|^{-d}$  is integrable, the result is exactly [11, Proposition 5.3.2]. The general case follows by monotone convergence. ■

The Radon property of  $m$  can now be proved.

**Theorem 1** *Let  $\alpha \in \mathbf{R}$  and define  $w(x) = |\det x|^\alpha$ . The function  $w$  is locally integrable on  $\mathbf{M}^d$  if and only if  $\alpha > -1$ .*

**Proof.** Let  $A \subset \mathbf{M}^d$  be relatively compact. Since  $\partial E$  is a nullset, we may assume that  $A \cap \partial E = \emptyset$ . Then there is a cube  $K \subset T(d)$ , say  $K = \prod_{i \leq j} I_{ij}$ , of bounded open intervals

$I_{ij}$  such that  $I_{ii} \subset (0, \infty)$  and  $I_{ij} \subset \mathbf{R}$  ( $i < j$ ), and such that  $A \subset O(d) \cdot K$ . Hence, the change-of-variable formula in Lemma 1 yields

$$\begin{aligned} \int_A |\det x|^\alpha dx &\leq \mu(O(d)) \left( \prod_{i < j} \int_{I_{ij}} dR_{ij} \right) \left( \int_K |\det R|^{\alpha+d} \prod_{i=1}^d R_{ii}^{-i} dR_{ii} \right) \\ &= \mu(O(d)) \left( \prod_{i < j} \int_{I_{ij}} dR_{ij} \right) \left( \prod_{i=1}^d \int_{I_{ii}} R_{ii}^{\alpha+d-i} dR_{ii} \right), \end{aligned}$$

where  $\mu$  is a non-normalized Haar measure on  $O(d)$ . The right side is finite, provided  $\alpha > -1$ . If on the other hand  $\alpha \leq -1$ , take  $I_{ij} = (0, 1)$  for all  $i \leq j$ , and set  $A = O(d) \cdot K$ . Then  $A$  is relatively compact, but  $\int_A w(x) dx = \infty$ . ■

Consider the following differential operator:

$$\nabla^* G = -(\nabla + (\delta - 1)x^{-\top}) \bullet G, \quad G \in C^1(E; \mathbf{M}^d).$$

Notice that  $\mathcal{A} = -\frac{1}{2}\nabla^*\nabla$ . With this in mind, the integration by parts formula below shows that  $\mathcal{A}$  is indeed a symmetric operator—indeed, taking  $G = \nabla g$  with  $g \in C_{c,\perp}^2$ , it implies that the equality

$$\langle f, -\mathcal{A}g \rangle = \frac{1}{2} \langle \nabla f, \nabla g \rangle$$

holds for all  $f \in C_c^1(E)$ . Here and in the sequel,  $\langle \cdot, \cdot \rangle$  denotes the inner product on  $L^2(E, m)$ . Specifically, we write

$$\langle f, g \rangle = \int_E f(x)g(x)m(dx) \quad \text{and} \quad \langle F, G \rangle = \int_E F(x) \bullet G(x)m(dx),$$

where  $f, g \in L^2(E, m)$  and  $F, G \in L^2(E, m; \mathbf{M}^d)$ . The overlapping notation should not cause any confusion.

**Theorem 2 (Integration by parts formula)** *Assume that  $\delta > 0$ , and let  $f \in C_c^1(E)$  and  $G \in C^1(E; \mathbf{M}^d)$ . If  $\delta \leq 1$ , assume that  $G(x)$  is tangent to  $\partial E$  at  $x$  for all  $x \in \partial E$ . If  $\delta < 1$ , assume in addition that  $G(x) \bullet x^{-\top}$  is locally bounded. Then*

$$\langle \nabla f, G \rangle = \langle f, -\nabla^* G \rangle.$$

**Proof.** See Section 5. ■

**Remark 3** It is not hard to show that local boundedness of  $G(x) \bullet x^{-\top}$  implies that  $G(x)$  is tangent to  $\partial E$  at  $x$  for all  $x \in \partial E$ .

The BESM process can now be defined as follows.

**Definition 1 (BESM semigroup)** *A strongly continuous sub-Markovian contraction semigroup  $(T_t : t \geq 0)$  on  $L^2(E, m)$  is called a BESM( $\delta, d$ ) semigroup if its generator extends  $(\mathcal{A}, C_{c,\perp}^2)$ .*

An  $E$ -valued Markov process  $X$  is called  $m$ -symmetric if its transition function  $p_t(x, dy)$  is  $m$ -symmetric. In this case the operators  $\int_E f(y)p_t(x, dy)$ , where  $f$  is bounded and in  $L^2(E, m)$ , can be extended to all of  $L^2(E, m)$ , see [12, page 30]. This extension is called the  $L^2(E, m)$  semigroup of  $X$ .

**Definition 2 (BESM process)** *An  $E$ -valued  $m$ -symmetric Markov process whose  $L^2(E, m)$  semigroup is a BESM( $\delta, d$ ) semigroup is called a BESM( $\delta, d$ ) process.*

A substantial portion of this paper is devoted to the question of uniqueness of the BESM semigroup and process. Existence, on the other hand, is straightforward via the theory of Dirichlet forms. In view of the discussion so far it is natural to introduce the bilinear symmetric form

$$\mathcal{E}(f, g) = \frac{1}{2} \langle \nabla f, \nabla g \rangle, \quad f, g \in C_c^1(E).$$

This form is closable on  $L^2(E, m)$ , as can be deduced from Theorem 2 as follows. Pick a sequence  $(f_n)$  in  $C_c^1(E)$  converging to zero in  $L^2(E, m)$ , such that  $\lim_{m,n} \mathcal{E}(f_n - f_m, f_n - f_m) = 0$ . We must show  $\lim_n \mathcal{E}(f_n, f_n) = 0$ . Since  $(\nabla f_n)$  is a Cauchy sequence in  $L^2(E, m; \mathbf{M}^d)$ , it has a strong limit  $F$ . For any  $\Phi \in C_c^\infty(E; \mathbf{M}^d)$  vanishing near  $\partial E$ , Theorem 2 yields

$$\langle F, \Phi \rangle = \lim_{n \rightarrow \infty} \langle \nabla f_n, \Phi \rangle = \lim_n \langle f_n, \nabla^* \Phi \rangle = 0.$$

It follows that  $F = 0$   $m$ -a.e., as required. We denote the closure of  $(\mathcal{E}, C_c^1(E))$  by

$$(\mathcal{E}, D(\mathcal{E})).$$

An application of [12, Theorem 3.1.2] then shows (after routine verification of the conditions of that theorem) that  $(\mathcal{E}, D(\mathcal{E}))$  is a Dirichlet form. Moreover, it is

(i) *regular*:  $D(\mathcal{E}) \cap C_c(E)$  is  $\|\cdot\|_\infty$ -dense in  $C_c(E)$  and  $\mathcal{E}_1$ -dense in  $D(\mathcal{E})$ , where

$$\mathcal{E}_1(f, f) = \langle f, f \rangle + \mathcal{E}(f, f).$$

(ii) *strongly local*: for every  $f, g \in D(\mathcal{E})$  with compact support such that  $g$  is constant on a neighborhood of the support of  $f$ , we have  $\mathcal{E}(f, g) = 0$ .

The integration by parts formula implies that the equality  $\mathcal{E}(f, g) = \langle -\mathcal{A}f, g \rangle$  holds for all  $f \in C_{c,\perp}^2$  (indeed, for all  $f \in C_c^2$  if  $\delta > 1$ ) and all  $g \in C_c^1(E)$ , hence for all  $g \in D(\mathcal{E})$  by continuity. It follows that the generator associated with  $\mathcal{E}$  coincides with  $\mathcal{A}$  when acting on functions  $f$  as above. With some abuse of notation, we therefore let

$$(\mathcal{A}, D(\mathcal{A}))$$

denote the generator of  $\mathcal{E}$ , noting that the domain  $D(\mathcal{A})$  contains  $C_{c,\perp}^2$ , and even contains  $C_c^2(E)$  if  $\delta > 1$ . In particular, the semigroup  $(T_t : t > 0)$  on  $L^2(E, m)$  associated with  $\mathcal{E}$  and  $\mathcal{A}$  is a BESM( $\delta, d$ ) semigroup.

A corresponding BESM( $\delta, d$ ) process is then obtained as the  $m$ -symmetric Hunt process  $X$  on  $E$  corresponding to the Dirichlet form  $(\mathcal{E}, D(\mathcal{E}))$ , see [12, Theorems 7.2.1 and 7.2.2].

The strongly local property of  $\mathcal{E}$  implies that this process has continuous paths. However, it is not guaranteed *a priori* that  $X$  is conservative; we now prove that it is, thereby obtaining existence of the BESM( $\delta, d$ ) process. In the following, let  $\mathbb{P}^x$  be the law of  $X$  starting from  $x \in E$ .

**Proposition 1** *The Dirichlet form  $\mathcal{E}$  is conservative, i.e. the semigroup  $(T_t : t > 0)$  satisfies  $T_t 1 = 1$  for all  $t > 0$ . Consequently,  $X$  can be chosen so that*

$$\mathbb{P}^x(X_t \in E \text{ for all } t \geq 0) = 1 \quad \text{for all } x \in E.$$

**Remark 4** For  $f \in L^\infty(E, m)$  with  $f \geq 0$ , one defines  $T_t f = \lim_n T_t f_n$ , where  $f_n \in L^2(E, m)$  and  $f_n \uparrow f$   $m$ -a.e. This definition is independent of the choice of sequence  $(f_n)$ .

**Proof.** By [12, Theorem 1.6.6],  $\mathcal{E}$  is conservative if there is a sequence  $(f_n) \subset D(\mathcal{E})$  such that  $0 \leq f_n \leq 1$  and  $\lim_n f_n = 1$   $m$ -a.e., and such that

$$\lim_n \mathcal{E}(f_n, g) \text{ holds for any } g \in D(\mathcal{E}) \cap L^1(E, m).$$

To construct such a sequence, let  $\phi \in C^2(\mathbf{R})$  satisfy  $\phi(t) = 1$  for  $t \leq 0$ ,  $\phi(t) = 0$  for  $t \geq 1$ , and  $|\phi'(t)|$  and  $|\phi''(t)|$  uniformly bounded. For  $n \geq 1$  define

$$f_n(x) = \phi(\|x\| - n).$$

Differentiating twice yields

$$\begin{aligned} \nabla f_n(x) &= \phi'(\|x\| - n) \frac{x}{\|x\|}, \\ \Delta f_n(x) &= \phi''(\|x\| - n) + \phi'(\|x\| - n) \frac{d^2 - 1}{\|x\|}. \end{aligned}$$

In particular we have  $f_n \in C_{c,\perp}^2 \subset D(\mathcal{A})$ . Since also  $\nabla f_n$  and  $\Delta f_n$  both vanish outside the set  $E_n = \{x \in E : n \leq \|x\| < n+1\}$ , we obtain

$$|\mathcal{E}(f_n, g)| = \left| \int_{E_n} \mathcal{A}f_n(x) g(x) m(dx) \right| \leq \text{ess sup}_{x \in E_n} |\mathcal{A}f_n(x)| \int_{E_n} |g(x)| m(dx),$$

where Hölder's inequality was applied. For  $n \geq 1$ , the essential supremum is bounded by a constant  $c > 0$  that is independent of  $n$ . Hence

$$\sum_{n \geq 1} |\mathcal{E}(f_n, g)| \leq c \sum_{n \geq 0} \int_{E_n} |g(x)| m(dx) = c \|g\|_{L^1(E, m)} < \infty.$$

We deduce that  $\lim_n |\mathcal{E}(f_n, g)| = 0$ , showing that  $\mathcal{E}$  is conservative. The statement about  $X$  now follows from [12, Exercise 4.5.1]. ■

We finish this section with a preliminary result regarding  $C_{c,\perp}^2$ . In Section 4, this set will be studied further.

**Proposition 2** *The set  $C_{c,\perp}^2$  is a core for  $\mathcal{E}$ .*

**Proof.** We apply a slightly modified version of [8, Theorem 5.1], where in the proof we use Theorem 6.2 instead of Theorem 2.3 (both in the same reference.) This shows that  $C_{c,\perp}^2$  is dense in  $C_c^1(E)$  with respect to both  $\|\cdot\|_\infty$  and  $\mathcal{E}_1$ . ■

### 3 Relation to the Wishart process

We now describe the properties of the transformed processes  $X^\top X$ ,  $\|X\|$ ,  $\det X$ , where  $X$  is a BESM( $\delta, d$ ) process realized as an  $m$ -symmetric Hunt process with BESM( $\delta, d$ ) semi-group, as discussed in the previous section. We always take  $\delta > 0$ . The main observation is the following. Let

$$\Phi : \mathbf{M}^d \rightarrow \mathbf{S}_+^d, \quad x \mapsto x^\top x,$$

and define

$$\mathcal{A}^{\text{WIS}} g(z) = \text{tr}(2z\nabla^2 + (d-1+\delta)\nabla)g(z), \quad z \in \mathbf{S}_+^d, \quad g \in C_c^\infty(\mathbf{S}_+^d).$$

This is the generator of the WIS( $\alpha, d$ ) process with  $\alpha = d-1+\delta$ , see [4]. Now, for any  $g \in C_c^\infty(\mathbf{S}_+^d)$  and any  $x \in E$ , one readily verifies the identities

$$\nabla(g \circ \Phi)(x) = 2x\nabla g(x^\top x), \quad \frac{1}{2}\Delta(g \circ \Phi)(x) = \text{tr}(2x^\top x\nabla^2 + d\nabla)g(x^\top x).$$

Consequently we have

$$g \circ \Phi \in D(\mathcal{A}) \quad \text{and} \quad \mathcal{A}(g \circ \Phi) = (\mathcal{A}^{\text{WIS}} g) \circ \Phi \in C_c^\infty(\mathbf{S}_+^d). \quad (2)$$

This relation will let us deduce that  $X^\top X$  is a Wishart process. The proof uses the following standard result.

**Lemma 2** *Pick any  $f \in D(\mathcal{A})$  such that  $\mathcal{A}f$  is locally  $m$ -integrable on  $E$ . For quasi-every  $x \in E$ , the process*

$$f(X_t) - f(x) - \int_0^t \mathcal{A}f(X_s)ds, \quad t \geq 0, \quad (3)$$

*is a square integrable martingale under  $\mathbb{P}^x$ , and  $\int_0^t |\mathcal{A}f(X_s)|ds < \infty$  for all  $t \geq 0$ ,  $\mathbb{P}^x$ -a.s.*

**Proof.** This can be deduced from [12, Theorem 5.2.2, Corollary 5.4.1, and Theorem 5.1.3]. The local integrability assumption is to ensure that  $|\mathcal{A}f(x)|m(dx)$  is a Radon measure, which is needed in [12, Corollary 5.4.1]. ■

**Remark 5** The exceptional set for which the conclusion of the lemma fails depends on the function  $f$  in general. However, while it will not be proved it here, we conjecture that  $X$  can be chosen so that its transition function is absolutely continuous with respect to  $m$ . This implies in particular that if  $f$  and  $g$  are measurable and satisfy  $f = g$ ,  $m$ -a.e., then the equality  $f(X_t) = g(X_t)$  holds  $\mathbb{P}^x$ -a.s. for every  $x \in E$  and every  $t \geq 0$ . By a standard argument using [10, Proposition 1.1.5], where the Banach space  $L$  is taken to be  $L^2(E, m)$ , one can then show that (3) defines a  $\mathbb{P}^x$  martingale for every  $f \in D(\mathcal{A})$  and every  $x \in E$  such that  $\mathbb{E}^x[\int_0^t |\mathcal{A}f(X_s)|ds] < \infty$ ,  $t \geq 0$ . In particular the “quasi-every” caveat can be removed in this case.

**Theorem 6** *Let  $X$  be a BESM( $\delta, d$ ) process as above.*

- (i)  $X^\top X$  is a WIS( $\alpha, d$ ) process, where  $\alpha = d-1+\delta$ .
- (ii)  $\|X\|$  is a BES( $\alpha$ ) process, where  $\alpha = d(d-1+\delta)$ .



(iii)  $\det X$  is a time-changed  $\text{BES}(\delta)$  process. More specifically, define

$$a(t) = \int_0^t \|\text{adj } X_s\|^2 ds, \quad \xi_u = \det X_{b(u)},$$

where  $b(u) = \inf\{t \geq 0 : a(t) \geq u\}$  is the left-continuous inverse of  $a$ . Then  $\xi$  is a  $\text{BES}(\delta)$  process stopped at  $a(\infty)$ .

**Remark 7** The precise meaning of part (i) of the theorem is that for all  $x$  outside some exceptional set  $N$ , the law of  $X^\top X$  under  $\mathbb{P}^x$  is that of a  $\text{WIS}(\alpha, d)$  process. The statements (ii) and (iii) should be interpreted analogously. Due to Remark 5, absolute continuity of the transition function would let us take  $N$  to be empty.

**Proof.** Lemma 2 together with (2) imply that for any countable family  $\mathcal{C} \subset C_c^\infty(\mathbb{S}_+^d)$  there is an exceptional set  $N \subset E$  such that (3) is a square integrable martingale under  $\mathbb{P}^x$  for all  $x \notin N$  and all  $f = g \circ \Phi$  with  $g \in \mathcal{C}$ . Setting  $Z_t = X_t^\top X_t$ , it follows that the process

$$g(Z_t) - g(x^\top x) - \int_0^t \mathcal{A}^{\text{WIS}} g(Z_s) dx, \quad t \geq 0,$$

is a martingale under  $\mathbb{P}^x$  for all  $g \in \mathcal{C}$ . Using standard arguments (see for instance [20, Theorem V.20.1]) this lets us deduce that  $X^\top X$  is a weak solution to the stochastic differential equation

$$dZ_t = \sqrt{Z_t} dW_t + dW_t^\top \sqrt{Z_t} + (d-1+\delta) \text{Id}t, \quad Z_0 = x^\top x,$$

for all  $x \notin N$ , where  $W$  is  $\mathbf{M}^d$ -valued Brownian motion under  $\mathbb{P}^x$ . (This also uses the well-known correspondence between the above stochastic differential equation and  $\mathcal{A}^{\text{WIS}}$ .) Hence (i) follows. Part (ii) is immediate from the well-known fact that the trace of a  $\text{WIS}(\alpha, d)$  process is a  $\text{BESQ}(\alpha d)$  process. For Part (iii) we set  $Z = X^\top X$  and note (see [4, Section 4]) that  $\det Z$  satisfies

$$\det Z_t = \det Z_0 + 2 \int_0^t \sqrt{\det Z_s} \sqrt{\text{tr}(\text{adj } Z_s)} d\beta_s + \delta \int_0^t \text{tr}(\text{adj } Z_s) ds$$

for some standard Brownian motion  $\beta$ . Defining

$$\tilde{\beta}_u = \int_0^{a^{-1}(u)} \sqrt{\text{tr}(\text{adj } Z_s)} d\beta_s,$$

which is Brownian motion on  $[0, a(\infty))$ , and using the identity  $\text{adj } z = (\text{adj } x)^\top \text{adj } x$ , we get

$$\det Z_t = \det Z_0 + 2 \int_0^t \sqrt{\det Z_s} d\tilde{\beta}_{a(s)} + \delta a(t).$$

Applying the time-change  $u = a(t)$  it follows that  $\det Z_{b(\cdot)}$  satisfies the stochastic differential equation for the  $\text{BESQ}(\delta)$  process, stopped at  $a(\infty)$  since  $b$  is constant on  $[a(\infty), \infty)$ . Since  $\det X_t = \sqrt{\det Z_t}$ , the result follows. ■

By a similar argument one can show that  $X$  itself is a weak solution to a certain stochastic differential equation when  $\delta > 1$ . Indeed, we saw in Section 2 that  $C_c^2(E) \subset D(\mathcal{A})$  if  $\delta > 1$ ,

and Theorem 1 implies that  $1/\det(x)$  is locally  $m$ -integrable in this case. We then deduce from Lemma 2 that (3) is a  $\mathbb{P}^x$  martingale for all  $f \in \mathcal{C} \subset C_c^2(E)$ ,  $\mathcal{C}$  countable, and all  $x$  outside some exceptional set  $N$ . This in turn implies that for all  $x \notin N$ ,  $X$  satisfies

$$X_t = x + W_t + \frac{\delta - 1}{2} \int_0^t X_s^{-\top} ds, \quad t \geq 0,$$

where  $W$  is Brownian motion under  $\mathbb{P}^x$ . In particular,  $X$  is a semimartingale. On the other hand, if  $0 < \delta < 1$ ,  $X$  is not a semimartingale—if it were, so would  $\det X$ , which however is a time-changed BES( $\delta$ ) process by Theorem 6 and hence not a semimartingale. In the case  $\delta = 1$  one can show that  $X$  is a semimartingale satisfying a Skorohod-type stochastic differential equation. We do not elaborate on the details here.

We close this section with a pathwise construction of the BESM process as the strong solution to a stochastic differential equation. This construction only works when  $\delta \geq 2$  and the process starts from  $E^o$ . We conjecture that the following result remains true for all  $\delta > 1$ .

**Proposition 3** *Suppose  $\delta \geq 2$ , and let  $W$  be standard  $\mathbf{M}^d$ -valued Brownian motion defined on some probability space. The stochastic differential equation with state space  $E^o$ ,*

$$dX_t = dW_t + \frac{\delta - 1}{2} X_s^{-\top} ds, \quad (4)$$

*has a unique strong solution.*

**Proof.** Since  $x \mapsto x^{-\top}$  is locally Lipschitz continuous on  $E^o$ , a strong solution  $X$  exists on  $[0, \zeta)$ , where  $\zeta = \lim_n \inf\{t \geq 0 : \det X_t < n^{-1} \text{ or } \|X_t\| > n\}$ . We claim that  $\zeta = \infty$ . To see this, set  $Z = X^\top X$  and note that we have

$$Z_t = \int_0^t X_s dW_s + \int_0^t dW_s^\top X_s^\top + (d - 1 + \delta)It, \quad t < \zeta,$$

where we used the equality  $dW_t dW_t = dIdt$ . Defining  $\widetilde{W}_t = \int_0^t (X_s^\top X_s)^{-1/2} X_s dW_s$ ,  $t < \zeta$ , we have

$$Z_t = \int_0^t \sqrt{Z_s} d\widetilde{W}_s + \int_0^t d\widetilde{W}_s^\top \sqrt{Z_s} + (d - 1 + \delta)It, \quad t < \zeta,$$

and after verifying that  $\widetilde{W}$  is again  $\mathbf{M}^d$ -valued Brownian motion on  $[0, \zeta)$ , it follows that  $Z$  is a WIS( $d - 1 + \delta, d$ ) process on  $[0, \zeta)$  (this calculation is of course closely related to the one in the beginning of Section 3.) Since  $\delta \geq 2$ , this implies that  $\det X_t = \sqrt{\det Z_t}$  stays strictly positive, and that  $\|X_t\|^2 = \text{tr } Z_t$  is nonexplosive. Hence  $\zeta = \infty$  as claimed, and the result is proved. ■

## 4 Uniqueness

The goal of this section is to establish uniqueness of the BESM semigroup. Specifically, we will prove that  $(\mathcal{A}, C_{c,\perp}^2)$  is *Markov unique*. This means that there is at most one (and

hence exactly one) symmetric sub-Markovian strongly continuous contraction semigroup on  $L^2(E, m)$  whose generator extends  $(\mathcal{A}, C_{c, \perp}^2)$ , see [9, Definition 1.1.2]. Since the Hunt process corresponding to such a semigroup is unique up to equivalence<sup>1</sup>, this form of uniqueness will hold for any realization of the BESM process as a Hunt process.

Note that  $m(\partial E) = 0$ . Therefore  $L^2(E, m)$  and  $L^2(E^\circ, m)$  can be identified, implying that it is enough to prove Markov uniqueness of  $(\mathcal{A}, C_{c, \perp}^2)$  as an operator on the latter space. To do this we will apply a general result by Eberle [9, Corollary 3.2] that relies on studying the relationship between various weighted Sobolev spaces, which we now introduce. To simplify notation we henceforth write

$$\Omega = E^\circ = \{x \in \mathbf{M}^d : \det x > 0\}.$$

Observe that  $1/(\det x)^{\delta-1}$  is locally integrable on  $\Omega$ . Hence by [17, Theorem 1.5],  $L^2(\Omega, m)$  is continuously imbedded in  $L^2(\Omega, dx)$ . In particular, every  $f \in L^2(\Omega, m)$  has a unique distributional gradient  $Df$ , and one can define the *weak Sobolev space*

$$W^{1,2}(\Omega, m) = \left\{ f \in L^2(\Omega, m) : Df \in L^2(\Omega, m; \mathbf{M}^d) \right\}.$$

Equipped with the norm

$$\|f\|_{W^{1,2}(\Omega, m)} = \left( \int_{\Omega} |f(x)|^2 m(dx) + \int_{\Omega} \|Df(x)\|^2 m(dx) \right)^{1/2},$$

$W^{1,2}(\Omega, m)$  becomes a Hilbert space [17, Theorem 1.11]. We also consider the following *strong Sobolev spaces* (here the word *completion* is always meant with respect to the norm  $\|\cdot\|_{W^{1,2}(\Omega, m)}$ ):

$$H^{1,2}(\Omega, m) = \text{completion of } C^\infty(\Omega)$$

$$H_{\perp}^{1,2}(\Omega, m) = \text{completion of } C_{c, \perp}^2$$

$$H_0^{1,2}(\Omega, m) = \text{completion of } C_c^\infty(\Omega).$$

These are all Hilbert spaces by construction, and we automatically have

$$H_0^{1,2}(\Omega, m) \subset H_{\perp}^{1,2}(\Omega, m) \subset H^{1,2}(\Omega, m) \subset W^{1,2}(\Omega, m). \quad (5)$$

In the present section we will expend significant effort to show that for any  $\delta > 0$ , the last two inclusions are equalities; and if  $\delta \geq 2$ , all three inclusions are equalities. This will lead to Markov uniqueness of the BESM semigroup. Specifically, we will prove the following:

**Theorem 8** *The following statements hold.*

- (i) *If  $0 < \delta < 2$ , then  $H_{\perp}^{1,2}(\Omega, m) = W^{1,2}(\Omega, m)$ .*

---

<sup>1</sup>Two symmetric Hunt processes are called equivalent if their transition functions coincide outside a *properly exceptional set*, see Section 4.1 in [12].

(ii) If  $\delta \geq 2$ , then  $H_0^{1,2}(\Omega, m) = W^{1,2}(\Omega, m)$ .

In either case,  $(A, C_{c,\perp}^2)$  is Markov unique.

The proof of this theorem requires some preparation. As in the one-dimensional situation, the case  $0 < \delta < 2$  is the most delicate. The following result, interesting in its own right, is crucial for handling this case. We let  $|A| = \int_A dx$  denote the Lebesgue measure of a measurable subset  $A \subset \mathbf{M}^d$ .

**Theorem 9 (Muckenhoupt property)** *Let  $\alpha \in \mathbf{R}$  and define  $w(x) = |\det x|^\alpha$ .*

(i) *If  $-1 < \alpha \leq 0$ , then  $w$  lies in the Muckenhoupt  $A_1$  class. That is, there is a constant  $C > 0$  depending only on  $d$  and  $\alpha$ , such that*

$$\frac{1}{|B|} \int_B w(x) dx \leq C \inf_{x \in B} w(x)$$

*for every ball  $B \subset \mathbf{M}^d$ .*

(ii) *If  $-1 < \alpha < p - 1$ ,  $p > 1$ , then  $w$  lies in the Muckenhoupt  $A_p$  class. That is, there is a constant  $C > 0$  depending only on  $d$ ,  $\alpha$  and  $p$ , such that*

$$\left( \frac{1}{|B|} \int_B w(x)^p dx \right) \left( \frac{1}{|B|} \int_B w(x)^{-1/(p-1)} dx \right)^{p-1} \leq C$$

*for every ball  $B \subset \mathbf{M}^d$ .*

**Proof.** The proof of part (i) is given in Section 6. Part (ii) follows from (i) via [21, Proposition IX.4.3]. ■

An application of this theorem with  $\alpha = \delta - 1$  shows that  $|\det x|^{\delta-1}$  is an  $A_2$ -weight if  $0 < \delta < 2$ . In particular, [16, Theorem 2.5] then shows that the last inclusion in (5) is in fact an equality.

The other key ingredient in the proof of Theorem 8 is to show that the sets  $M_k$  of low-rank matrices are small. To measure the size of a set we use the following notion of capacity (referred to as the  $(1, m)$ -Sobolev capacity in [13]). For any subset  $A \subset \mathbf{M}^d$ , define

$$\text{Cap}(A) = \inf_f \int_{\mathbf{M}^d} (|f(x)|^2 + \|Df(x)\|^2) m(dx),$$

where the infimum is taken over all  $f \in H^{1,2}(\mathbf{M}^d, m)$  with  $f = 1$  on an open neighborhood of  $A$ , see [13, Definition 2.35]. We have the following result:

**Theorem 10** *Let  $\delta > 0$ . For  $k \in \{0, \dots, d-2\}$ , we have  $\text{Cap}(M_k) = 0$ . If  $\delta \geq 2$ , the same thing also holds for  $k = d-1$ .*

The core of the proof of Theorem 10 is an application of the following lemma, which bounds the growth of the determinant function near a point  $x \in M_k$ .

**Lemma 3** *Let  $k \leq d-1$ . There is a locally Lipschitz function  $c_k : M_k \rightarrow \mathbf{R}_+$  such that*

$$|\det(x+v)| \leq c_k(x) \|v\|^{d-k}, \quad x \in M_k, \ v \in \mathbf{M}^d, \ \|v\| \leq 1.$$

**Proof.** By [3, Corollary 5],

$$|\det(x + v) - \det(x)| \leq \sum_{i=1}^d p_{d-i}(\sigma_1(x), \dots, \sigma_d(x)) \|v\|^i,$$

where  $\sigma(x) = (\sigma_1(x), \dots, \sigma_d(x))$  is the vector of singular values of  $x$ , and  $p_i$  is the  $i$ :th elementary symmetric polynomial in  $d$  variables. Now,  $p_{d-i}(\sigma_1(x), \dots, \sigma_d(x))$  consists of a sum of terms, each of which is the product of  $d-i$  distinct elements of  $\sigma(x)$ . However, since  $\text{rank } x = k$ , only  $k$  of those elements are nonzero. Therefore the product must contain at least one zero factor whenever  $d-i > k$ , implying that  $p_{d-i}(\sigma_1(x), \dots, \sigma_d(x)) = 0$  for these  $i$ . Since in addition  $\|v\| \leq 1$  and  $\det x = 0$ , we get

$$|\det(x + v)| \leq \|v\|^{d-k} \sum_{i=d-k}^d p_{d-i}(\sigma_1(x), \dots, \sigma_d(x)).$$

The local Lipschitz property follows from the smoothness of  $p_{d-i}$  and the fact that the singular value map is Lipschitz continuous, see [14, Theorem 7.4.51]. ■

In proving Theorem 10, the case  $k = d - 1$ ,  $\delta = 2$  turns out to require separate treatment using the following lemma.

**Lemma 4** *For each  $\varepsilon < 1$  there is a Lipschitz function  $\phi_\varepsilon : \mathbf{R}_+ \rightarrow \mathbf{R}$  such that  $0 \leq \phi_\varepsilon \leq 1$ ,  $\phi_\varepsilon = 0$  on  $[\varepsilon, \infty)$ ,  $\phi_\varepsilon = 1$  on a neighborhood of zero, and*

$$\lim_{\varepsilon \downarrow 0} \int_{\mathbf{R}_+} |\phi'_\varepsilon(t)|^2 t \, dt = 0. \quad (6)$$

**Proof.** Define functions  $g_\varepsilon$  and  $h_\varepsilon$  on  $\mathbf{R}_+$  by

$$g_\varepsilon(t) = \left(1 - \left(\frac{t}{\varepsilon}\right)^\varepsilon\right)_+ \quad \text{and} \quad h_\varepsilon(t) = \begin{cases} (t/\varepsilon)^\varepsilon & t \in [0, \varepsilon^{1+1/\varepsilon}) \\ 2\varepsilon - \varepsilon^{-1/\varepsilon}t & t \in [\varepsilon^{1+1/\varepsilon}, 2\varepsilon^{1+1/\varepsilon}) \\ 0 & t \in [2\varepsilon^{1+1/\varepsilon}, \infty) \end{cases}$$

We claim that the function  $\phi_\varepsilon = g_\varepsilon + h_\varepsilon$  has the stated properties. It is not hard to check that  $0 \leq \phi_\varepsilon \leq 1$  and that  $\phi_\varepsilon$  equals zero on  $[\varepsilon, \infty)$  and one on  $[0, \varepsilon^{1+1/\varepsilon})$ . The Lipschitz property then follows easily. It remains to verify (6). First, note that

$$\int_{\mathbf{R}_+} |g'_\varepsilon(t)|^2 t \, dt = \varepsilon^{2-2\varepsilon} \int_0^\varepsilon t^{2\varepsilon-1} dt = \frac{\varepsilon}{2} \rightarrow 0 \quad (\varepsilon \downarrow 0).$$

Moreover, since  $|h'_\varepsilon| = |g'_\varepsilon|$  on  $[0, \varepsilon^{1+1/\varepsilon})$ , and since

$$\int_{\varepsilon^{1+1/\varepsilon}}^{2\varepsilon^{1+1/\varepsilon}} |h'_\varepsilon(t)|^2 t \, dt = \varepsilon^{-2/\varepsilon} \frac{(2\varepsilon^{1+1/\varepsilon})^2 - (\varepsilon^{1+1/\varepsilon})^2}{2} = \frac{3}{2}\varepsilon^2 \rightarrow 0 \quad (\varepsilon \downarrow 0),$$

it follows that  $\lim_{\varepsilon \downarrow 0} \int_{\mathbf{R}_+} |h'_\varepsilon(t)|^2 t \, dt = 0$ . We now deduce (6). ■

The proof of Theorem 10 exploits the fact that  $M_k$  is a smooth manifold—specifically, we need the notion of an *open tube segment*. To introduce this notion, let  $M$  be any smooth  $n_1$ -dimensional embedded submanifold of  $\mathbf{R}^n$  ( $n_1 < n$ ), and fix any point  $\bar{x} \in M$ . Then take an open neighborhood  $U \subset \mathbf{R}^n$  of  $\bar{x}$ , and a diffeomorphism

$$\Phi : U \rightarrow \Phi(U)$$

such that  $\Phi(M \cap U) = \widetilde{M} \times \{0\} \in \mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$ ,  $n_2 = n - n_1$ , for some open subset  $\widetilde{M} \subset \mathbf{R}^{n_1}$  containing  $\Phi(\bar{x})$  (these objects exist by definition of smooth manifold). Let  $G$  be an open ball in  $\mathbf{R}^{n_1}$ , centered at  $\Phi(\bar{x})$ , with its closure contained in  $\widetilde{M}$ . Let  $B_\varepsilon$  be the open ball in  $\mathbf{R}^{n_2}$ , centered at the origin, with radius  $\varepsilon$ . For all sufficiently small  $\varepsilon$ , say  $\varepsilon \leq \varepsilon_0$ , the closure of  $G \times B_\varepsilon$  lies in  $\Phi(U)$ . Define the set

$$U_\varepsilon = \Phi^{-1}(G \times B_\varepsilon).$$

We refer to  $U_\varepsilon$  as an *open tube segment* around  $M$  centered at  $\bar{x}$ . It follows from the above that for each  $\varepsilon \leq \varepsilon_0$ ,  $U_\varepsilon$  is an open subset of  $\mathbf{R}^n$  containing  $\Phi^{-1}(G \times \{0\})$ . In our setting,  $\mathbf{R}^n$  is identified with  $\mathbf{M}^d$ , so we have  $n = d^2$ , and  $M$  is identified with  $M_k$  for the particular  $k$  under consideration, so that  $n_2 = (d - k)^2$  and  $n_1 = d^2 - (d - k)^2$ .

We now have all the ingredients needed for the proof of Theorem 10.

**Proof of Theorem 10.** Let  $U_\varepsilon$  be an open tube segment around  $M_k$  centered at some fixed  $\bar{x}$ , as described above. Let  $\pi : \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \rightarrow \mathbf{R}^{n_2}$  be the coordinate projection onto the last  $n_2$  coordinates. Let  $\varphi \in C^\infty([0, \infty))$  be a cutoff function valued in  $[0, 1]$ , equal to one on  $[0, 1/2]$ , equal to zero on  $[1, \infty)$ , and with  $|\varphi'(t)| \leq 3$  for all  $t \geq 0$ . Define

$$\widetilde{g}_\varepsilon(\widetilde{x}) = \varphi(\varepsilon^{-1} \|\pi(\widetilde{x})\|), \quad \widetilde{x} \in G \times B_{\varepsilon_0},$$

and then

$$g_\varepsilon(x) = \widetilde{g}_\varepsilon \circ \Phi(x), \quad x \in U_{\varepsilon_0}. \quad (7)$$

It is clear that  $g_\varepsilon$  lies in  $C^\infty(U_{\varepsilon_0})$ , vanishes outside  $U_\varepsilon$ , and is equal to one on an open neighborhood  $U'_\varepsilon$  of  $\Phi^{-1}(G \times \{0\})$ . Let us compute  $\nabla g_\varepsilon$ . Pick any  $v \in \mathbf{M}^d$  and any  $x \in U_{\varepsilon_0}$ . By the chain rule,

$$\nabla g_\varepsilon(x) \bullet v = \nabla \widetilde{g}_\varepsilon(\Phi(x)) \bullet \Phi'(x)[v],$$

where  $\Phi'(x)[v]$  denotes the Fréchet derivative of  $\Phi$  at  $x$  applied to  $v$ . The chain rule also gives, for  $\widetilde{x} \in G \times B_{\varepsilon_0}$ ,

$$\nabla \widetilde{g}_\varepsilon(\widetilde{x}) \bullet v = \frac{1}{\varepsilon} \varphi'(\varepsilon^{-1} \|\pi(\widetilde{x})\|) \frac{\pi(\widetilde{x})}{\|\pi(\widetilde{x})\|} \bullet \pi'(\widetilde{x})[v].$$

The derivative of  $\pi$  is given by  $\pi'(\widetilde{x})[v] = \pi(v)$ , and since orthogonal projections are contractions, we deduce

$$\|\nabla \widetilde{g}_\varepsilon(\widetilde{x})\| \leq \frac{3}{\varepsilon},$$

and consequently

$$\|\nabla g_\varepsilon(x)\| \leq \frac{3}{\varepsilon}C,$$

where  $C = \sup_{x \in U_{\varepsilon_0}} \|\Phi'(x)\|$  (operator norm) is finite due to the fact that the (compact) closure of  $U_{\varepsilon_0}$  is contained in a set where  $\Phi$  is smooth. It follows that

$$\int_{U_{\varepsilon_0}} (|g_\varepsilon(x)|^2 + \|\nabla g_\varepsilon(x)\|^2) m(dx) \leq \left(1 + \frac{9C^2}{\varepsilon^2}\right) m(U_\varepsilon). \quad (8)$$

Suppose we can show that the right side tends to zero as  $\varepsilon \downarrow 0$ . Then, multiplying  $g_\varepsilon$  by a smooth function  $\psi$  that is zero outside  $U_{\varepsilon_0}$  and one on some open neighborhood  $U'_{\varepsilon_0}$  of  $\bar{x}$ , we obtain functions  $f_\varepsilon \in H^{1,2}(\mathbf{M}^d, m)$  certifying that  $\text{Cap}(M_k \cap U'_{\varepsilon_0}) = 0$ . Since  $M_k$  is covered by a countable union of such sets  $U'_{\varepsilon_0}$ , [13, Theorem 2.37] implies that  $\text{Cap}(M_k) = 0$ , as required.

Thus, it remains to show that  $\varepsilon^{-2}m(U_\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . A change of variables yields

$$m(U_\varepsilon) = \int_{U_\varepsilon} |\det x|^{\delta-1} dx = \int_{G \times B_\varepsilon} |\det \Phi^{-1}(y, v)|^{\delta-1} J(y, v) dy \otimes dv,$$

where  $J(y, v)$  is the Jacobian determinant, and where we write  $\mathbf{R}^n = \mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$ ,  $\tilde{x} = (y, v)$ ,  $d\tilde{x} = dy \otimes dv$  (Lebesgue measures). Since  $\Phi^{-1}$  is smooth on a set containing the closure of  $G \times B_\varepsilon$ , it is Lipschitz on  $G \times B_\varepsilon$  with some constant  $\kappa_1$ . Moreover,  $J(y, v)$  is bounded by some constant  $\kappa_2$ . Together with Lemma 3 (and the fact that  $\Phi^{-1}(y, 0) \in M_k$ ), this yields

$$\begin{aligned} m(U_\varepsilon) &\leq \kappa_2 \int_{G \times B_\varepsilon} |\det (\Phi^{-1}(y, 0) + \Phi^{-1}(y, v) - \Phi^{-1}(y, 0))|^{\delta-1} dy dv \\ &\leq \kappa_2 \int_{G \times B_\varepsilon} c_k \circ \Phi^{-1}(y, 0) \|\Phi^{-1}(y, v) - \Phi^{-1}(y, 0)\|^{(d-k)(\delta-1)} dy dv \\ &\leq \kappa_1 \kappa_2 \int_{G \times B_\varepsilon} c_k \circ \Phi^{-1}(y, 0) \|v\|^{(d-k)(\delta-1)} dy dv \\ &= \kappa_1 \kappa_2 \int_G c_k \circ \Phi^{-1}(y, 0) dy \int_{B_\varepsilon} \|v\|^{(d-k)(\delta-1)} dv, \end{aligned}$$

where  $c_k(\cdot)$  is as in Lemma 3. The integral over  $G$  is finite by (uniform) continuity of  $c_k$  and  $\Phi^{-1}$ , so we get

$$m(U_\varepsilon) \leq C \int_{B_\varepsilon} \|v\|^{(d-k)(\delta-1)} dv, \quad \varepsilon \leq \varepsilon_0 \wedge 1,$$

for some constant  $C > 0$  that does not depend on  $\varepsilon$ . Since  $B_\varepsilon = \varepsilon B_1$ , the right side equals

$$C \int_{B_1} \|\varepsilon v\|^{(d-k)(\delta-1)} \varepsilon^{n_2} dv = \varepsilon^{-(d-k)(1-\delta)+n_2} C \int_{B_1} \|v\|^{-(d-k)(1-\delta)} dv.$$

Since the integral is over  $n_2$ -dimensional space, the right side is finite provided

$$(d-k)(1-\delta) < n_2 - 2. \quad (9)$$

But  $n_2 = (d - k)^2$ , so (9) is equivalent to  $(d - k)(d - k - 1 + \delta) > 2$ , which holds for all  $k \leq d - 2$  since  $\delta > 0$ . We conclude that there is a constant  $C' > 0$ , independent of  $\varepsilon$ , such that

$$\frac{1}{\varepsilon^2} m(U_\varepsilon) \leq C' \varepsilon^{-2-(d-k)(1-\delta)+n_2}.$$

Since, as we just saw, (9) holds, this quantity tends to zero as  $\varepsilon$  tends to zero. This finishes the proof of the assertion for  $\delta > 0$ .

If  $\delta > 2$ , then (9) holds also for  $k = d - 1$ , which takes care of this case as well. The only case that remains to prove is  $\delta = 2$  and  $k = d - 1$ . The argument is exactly analogous to the one above, except that  $\tilde{g}_\varepsilon$  is now given by

$$\tilde{g}_\varepsilon(\tilde{x}) = \phi_\varepsilon(|\pi(\tilde{x})|),$$

where  $\phi_\varepsilon$  is the function from Lemma 4. Note that  $n_2 = (d - k)^2 = 1$ , so that  $\pi(\tilde{x})$  is a real number. The function  $g_\varepsilon$  is again given by (7). Next, instead of (8) we need a more precise estimate. Specifically, we have the inequality

$$\|\nabla g_\varepsilon(x)\| \leq C |\phi'_\varepsilon(|\pi \circ \Phi(x)|)|,$$

where as before,  $C = \sup_{x \in U_{\varepsilon_0}} \|\Phi'(x)\|$ . Hence, by the same calculations as above involving Lemma 3, we get

$$\begin{aligned} \int_{U_{\varepsilon_0}} \|\nabla g_\varepsilon(x)\|^2 m(dx) &\leq C^2 \int_{U_\varepsilon} |\phi'_\varepsilon(|\pi \circ \Phi(x)|)|^2 |\det x| dx \\ &= C^2 \int_{G \times B_\varepsilon} |\phi'_\varepsilon(|v|)|^2 |\det \Phi^{-1}(y, v)| J(y, v) dy \otimes dv \\ &\leq \kappa_1 \kappa_2 C^2 \int_G c_{d-1} \circ \Phi^{-1}(y, 0) dy \int_{-\varepsilon}^\varepsilon |\phi'_\varepsilon(|v|)|^2 |v| dv. \end{aligned}$$

By the property (6) of  $\phi_\varepsilon$  given in Lemma 4, the right side tends to zero as  $\varepsilon \downarrow 0$ . This concludes the proof. ■

Equipped with Theorems 9 and 10, we can make further progress toward proving Theorem 8. Let us start by briefly reviewing some properties of the space  $W^{1,2}(\Omega, m)$ . For any open set  $\Omega'$  with compact closure contained in  $\Omega$ , we have  $C^{-1} \leq (\det x)^{\delta-1} \leq C$  for some constant  $C > 1$  and all  $x \in \Omega'$ . Hence

$$\|\cdot\|_{W^{1,2}(\Omega', m)} \text{ and } \|\cdot\|_{W^{1,2}(\Omega', dx)} \text{ are equivalent,} \quad (10)$$

and the unweighted space  $W^{1,2}(\Omega', dx)$  coincides with  $W^{1,2}(\Omega', m)$ . One consequence is that any  $f \in W^{1,2}(\Omega, m)$  supported on a compact set  $K \subset \Omega$  can be approximated by functions from  $C_c^\infty(\Omega)$  via mollification: take  $\Omega'$  bounded with  $K \subset \Omega' \subset \overline{\Omega'} \subset \Omega$ , so that  $f \in W^{1,2}(\Omega', dx)$ , and use standard mollification results for this space.

Another useful consequence of (10) is that  $W^{1,2}(\Omega, m)$  is stable under truncation: If  $f \in W^{1,2}(\Omega, m)$  then  $g = f \wedge 1$  has a weak derivative with  $Dg = Df \mathbf{1}_{\{f < 1\}}$  (in which case, of course,  $\|g\|_{W^{1,2}(\Omega, m)} \leq \|f\|_{W^{1,2}(\Omega, m)}$ ). To see this, pick any test function  $\varphi \in C_c^\infty(\Omega)$



with support  $K \subset \Omega$ . Now choose  $\Omega'$  as above, and apply known truncation results for  $W^{1,2}(\Omega', dx)$  to get

$$\int_{\Omega'} (f(x) \wedge 1) \varphi(x) dx = - \int_{\Omega'} (Df(x) \mathbf{1}_{\{f(x) < 1\}}) \bullet \nabla \varphi(x) dx.$$

Since  $\varphi$  vanishes outside  $\Omega'$ , the above relation holds with  $\Omega$  replacing  $\Omega'$ . This proves the claim. It is then easy to show that

$$\text{the bounded elements of } W^{1,2}(\Omega, m) \text{ are dense.} \quad (11)$$

After this digression we continue with two technical lemmas. The first shows that any element of  $W^{1,2}(\Omega, m)$  is close to some element whose support is bounded away from  $\cup_{k \leq d-2} M_k$  (and from  $\partial\Omega = \cup_{k \leq d-1} M_k$  in case  $\delta \geq 2$ .)

**Lemma 5** *Let  $\delta > 0$  and pick  $f \in W^{1,2}(\Omega, m)$ . For any  $\varepsilon > 0$  there is  $g \in W^{1,2}(\Omega, m)$  with bounded support such that  $\|f - g\|_{W^{1,2}(\Omega, m)} < \varepsilon$  and  $g = 0$  on  $U \cap \Omega$  for some open set  $U \subset \mathbf{M}^d$  containing  $\cup_{k \leq d-2} M_k$ . If  $\delta \geq 2$ ,  $U$  can be chosen to contain  $\partial\Omega$ .*

**Proof.** It is enough to consider positive  $f$  with bounded support, and by (11) we may assume  $f$  is bounded. By Theorem 10 there are open neighborhoods  $U_n$  of  $\cup_{k \leq d-2} M_k$  (respectively of  $\partial\Omega$  if  $\delta \geq 2$ ) and elements  $h_n \in W^{1,2}(\mathbf{M}^d, m)$  such that  $h_n = 1$  on  $U_n$  and  $h_n \rightarrow 0$  in  $W^{1,2}(\Omega, m)$ . Replacing  $h_n$  by  $(h_n \wedge 1) \vee 0$  if necessary, we may assume  $0 \leq h_n \leq 1$ . Now define

$$f_n = (1 - h_n)f.$$

Then

$$\begin{aligned} \|f_n - f\|_{W^{1,2}(\Omega, m)}^2 &= \int_{\Omega} (|h_n(x)f(x)|^2 + \|h_n(x)Df(x) + f(x)Dh_n(x)\|^2) m(dx) \\ &\leq 2\|f\|_{\infty}^2 \|h_n\|_{W^{1,2}(\Omega, m)}^2 + 2 \int_{\Omega} |h_n(x)|^2 \|Df(x)\|^2 m(dx). \end{aligned}$$

Fix any  $\varepsilon > 0$ . Let  $A_{\kappa} = \{x \in E : \|Df(x)\| \leq \kappa\}$ , and pick  $\kappa$  large enough that  $\int_{A_{\kappa}^c} \|Df(x)\|^2 m(dx) < \varepsilon/2$ . With this  $\kappa$ ,

$$\begin{aligned} \int_{\Omega} |h_n(x)|^2 \|Df(x)\|^2 m(dx) &\leq \int_{A_{\kappa}} |h_n(x)|^2 \|Df(x)\|^2 m(dx) + \int_{A_{\kappa}^c} \|Df(x)\|^2 m(dx) \\ &< \kappa^2 \int_{\Omega} |h_n(x)|^2 m(dx) + \varepsilon/2. \end{aligned}$$

It follows that for  $n$  sufficiently large,

$$\|f_n - f\|_{W^{1,2}(\Omega, m)}^2 < \varepsilon.$$

We may now take  $g = f_n$ . ■

The proof of the next lemma uses a standard partition of unity argument (mimicking the proof of [1, Theorem 3.22]) together with a powerful extension theorem due to Chua [5]

for weighted Sobolev spaces with Muckenhoupt weights. In particular, therefore, we will rely on the Muckenhoupt property of  $|\det x|^{\delta-1}$  for  $0 < \delta < 2$ . One delicate point is to choose the covering sets so that Chua's theorem becomes applicable. Specifically, their intersections with  $\Omega$  must be so-called  $(\varepsilon, \delta)$ -domains. Here the tube segments discussed previously will again be useful.

**Lemma 6** *Let  $0 < \delta < 2$  and define  $\tilde{\Omega} = \Omega \setminus \{x \in \mathbf{M}^d : \text{rank } x \leq d-2\}$ . Then*

$$C_c^\infty(\tilde{\Omega}) \text{ is dense in } W^{1,2}(\Omega, m).$$

**Proof.** By Lemma 5 it suffices to approximate elements  $f \in W^{1,2}(\Omega, m)$  with bounded support and with  $f = 0$  on  $U \cap \Omega$  for an open subset  $U \subset \mathbf{M}^d$  containing  $\cup_{k \leq d-2} M_k$ . Define

$$K = \text{closure in } \mathbf{M}^d \text{ of } \{x \in \Omega : g(x) \neq 0\},$$

and note that  $K$  is compact. Also, by shrinking  $U$  if necessary we may assume that  $U \cap K = \emptyset$ . Now, for every  $x \in M_{d-1}$ , let  $U_x$  be an open tube segment around  $x$ , thin enough so that  $\text{rank } y \geq d-1$  for all  $y \in U_x$ . Note that the set

$$F = K \setminus \cup_{x \in M_{d-1}} U_x$$

is compact in  $M^d$  and is contained in  $\Omega$ . Hence there is an open set  $U_0$  such that  $F \subset U_0 \subset \Omega$ . Next, since  $K$  is compact, we may choose finitely many of the  $U_x$ , say  $U_1, \dots, U_n$ , such that  $K \subset \cup_{i=0}^n U_i$ . There are also open sets  $V_0, \dots, V_n$  with  $\bar{V}_i \subset U_i$  for each  $i$  and such that  $K \subset \cup_{i=0}^n V_i$ . Let  $\{\psi_i : i = 0, \dots, n\}$  be a smooth partition of unity subordinate to  $V_0, \dots, V_n$ , and define  $f_i = \psi_i f$ . Fix any  $\varepsilon > 0$  and suppose we can find  $\phi_i \in C_c^\infty(\tilde{\Omega})$  such that  $\|f_i - \phi_i\|_{W^{1,2}(\Omega, m)} < \varepsilon/(1+n)$ . Then, letting  $\phi = \sum_i \phi_i$ , we have  $\phi \in C_c^\infty(\tilde{\Omega})$  with

$$\|f - \phi\|_{W^{1,2}(\Omega, m)} \leq \sum_{i=0}^n \|f_i - \phi_i\|_{W^{1,2}(\Omega, m)} < \varepsilon,$$

as desired. We thus focus on finding suitable  $\phi_i$ . For  $i = 0$  one can simply take a mollification of  $f_0$ , thanks to (10). For  $i \in \{1, \dots, n\}$ , we define  $\Omega_i = \Omega \cap U_i$ . Then  $f_i \in W^{1,2}(\Omega_i, m)$ , and it is readily verified that  $\Omega_i$  is a Lipschitz domain. Indeed, by construction it is the image of a cylinder—the product of a Euclidean ball and a (short) open line segment—under a bi-Lipschitz map. Since  $|\det x|^{\delta-1}$  is a Muckenhoupt  $A_2$  weight by Theorem 9 (and since every Lipschitz domain is an  $(\varepsilon, \delta)$ -domain, see [15, p. 73]), the extension theorem of Chua [5, Theorem 1.1] yields an element  $g_i \in W^{1,2}(\mathbf{M}^d, m)$  with  $g_i = f_i$  on  $\Omega_i$ . Since  $f_i = 0$  outside  $V_i$ , we may multiply  $g_i$  by a smooth cutoff function that is one on  $V_i$  and zero on  $U_i^c$ , and thus assume that  $g_i = f_i$  on all of  $\Omega$ . The desired  $\phi_i \in C_c^\infty(\tilde{\Omega})$  is now obtained by mollification of  $g_i$ . Indeed, the support of  $\phi_i$  can be made bounded away from  $\cup_{k \leq d-2} M_k$ , and we have

$$\|\phi_i - f_i\|_{W^{1,2}(E, m)} = \|\phi_i - g_i\|_{W^{1,2}(E, m)} \leq \|\phi_i - g_i\|_{W^{1,2}(\mathbf{M}^d, m)},$$

where the right side can be made arbitrarily small. ■

We are now finally ready to prove our main result leading to Markov uniqueness of the BESM semigroup.

**Proof of Theorem 8.** In view of (5), it is clear that in order to prove (i) (respectively (ii)) it suffices to prove  $W^{1,2}(\Omega, m) \subset H_{\perp}^{1,2}(\Omega, m)$  (respectively  $W^{1,2}(\Omega, m) \subset H_0^{1,2}(\Omega, m)$ ).

We start with part (ii), and pick  $f \in W^{1,2}(\Omega, m)$ . By Lemma 5,  $f$  can be approximated by some  $g \in W^{1,2}(\Omega, m)$  whose support is bounded and bounded away from  $\partial\Omega$ . By mollification (relying on (10)),  $g$  can in turn be approximated by  $h \in C_c^\infty(\Omega)$ . Since  $H_0^{1,2}(\Omega, m)$  is the completion of the set of such functions, we obtain  $f \in H_0^{1,2}(\Omega, m)$ , as desired.

Now consider part (i). By Lemma 6 the result follows if any given  $f \in C_c^\infty(\tilde{\Omega})$  can be approximated in  $W^{1,2}(\Omega, m)$  by  $g \in C_{c,\perp}^2$ . For this we can apply the same modified version of [8, Theorem 5.1] as in the proof of Proposition 2 above. However, since  $f$  is zero in a neighborhood of  $\cup_{k \leq d-2} M_k$ , there is a simple direct argument that we now give in order to make the presentation self-contained.

For each  $\varepsilon > 0$ , let  $\phi_\varepsilon \in C^2(\mathbb{R}_+)$  satisfy the following properties:

- (i)  $\phi_\varepsilon(t) = 0$  for  $t \geq \varepsilon$ ,
- (ii)  $\frac{1-\phi'_\varepsilon(t)}{t}$  and  $\frac{\phi_\varepsilon(t)}{t}$  are bounded in  $t$  (where the bound may depend on  $\varepsilon$ ),
- (iii)  $|\phi_\varepsilon(t)| \leq \varepsilon$  and  $|\phi'_\varepsilon(t)| \leq 2$  for all  $t > 0$ .

This is clearly possible when  $\varepsilon = 1$ , and hence also in general by setting  $\phi_\varepsilon(t) = \varepsilon\phi_1(t/\varepsilon)$ . Let  $K$  denote the support of  $f$ . With  $g(x) = \frac{\nabla f(x) \bullet \nabla \det(x)}{\|\nabla \det(x)\|^2}$  we have

$$\nabla f(x) = g(x) \nabla \det(x) + G(x), \quad (12)$$

where  $G(x) \bullet \nabla \det(x) = 0$  for all  $x$ . Note that  $\nabla \det(x) = \text{adj } x^\top \neq 0$  for all  $x \in K$ , so that  $g$  is well-defined if we set it to zero outside  $K$ . In fact,  $g \in C_c^\infty(E)$ . For  $\varepsilon > 0$ , consider the function

$$h_\varepsilon(x) = f(x) - \phi_\varepsilon(\det x)g(x).$$

Straightforward calculations using the chain and product rules together with (12) yield the following two expressions for  $\nabla h_\varepsilon$ :

$$\begin{aligned} \nabla h_\varepsilon(x) &= \nabla f(x) - \phi'_\varepsilon(\det x)g(x)\nabla \det(x) - \phi_\varepsilon(\det x)\nabla g(x) \\ &= (1 - \phi'_\varepsilon(\det x))\nabla f(x) + \phi'_\varepsilon(\det x)G(x) - \phi_\varepsilon(\det x)\nabla g(x). \end{aligned}$$

We first show that  $h_\varepsilon$  approximates  $f$  in  $W^{1,2}(\Omega, m)$  if  $\varepsilon$  is chosen sufficiently small. But this is an immediate consequence of the inequalities

$$|f(x) - h_\varepsilon(x)| \leq \varepsilon |g(x)|$$

and

$$\|\nabla f(x) - \nabla h_\varepsilon(x)\| \leq \left(2|g(x)|\|\nabla \det(x)\| + \|\nabla g(x)\|\right) \mathbf{1}_{[0,\varepsilon]}(\det x),$$

together with the fact that  $m(\{x \in K : 0 \leq \det x \leq \varepsilon\})$  tends to zero as  $\varepsilon \downarrow 0$ . It thus suffices to check that  $h_\varepsilon \in C_{c,\perp}^2$ . Clearly  $h_\varepsilon \in C_c^2(\Omega)$ . Moreover, the second of the above expressions for  $\nabla h_\varepsilon$  together with the orthogonality of  $\nabla \det(x)$  and  $G(x)$ , as well as the fact that  $x^{-\top} = \nabla \det(x) / \det(x)$ , yield

$$x^{-\top} \bullet \nabla h_\varepsilon(x) = \frac{1 - \phi'_\varepsilon(\det x)}{\phi_\varepsilon(\det x)} \nabla \det(x) \bullet \nabla f(x) - \frac{\phi_\varepsilon(\det x)}{\det x} \nabla \det(x) \bullet \nabla g(x).$$

The properties of  $\phi_\varepsilon$  imply that the right side is bounded, as required. Part (ii) of the theorem is thus proved.

It remains to prove Markov uniqueness. But this follows directly from the basic criterion for Markov uniqueness given in [9, Corollary 3.2], since for any  $\delta > 0$  we have  $W^{1,2}(\Omega, m) = H_{\perp}^{1,2}(\Omega, m)$ . ■

**Remark 11** Using results in [8], the space  $C_{c,\perp}^2$  is dense in  $C_c^2(E)$  with respect to the norm  $\|\cdot\|_{W^{1,2}(\Omega, m)}$ . An alternative approach to proving Theorem 8 (i) would therefore be to show directly that  $C_c^2(E)$  is dense in  $W^{1,2}(\Omega, m)$ , for example by showing that  $\Omega$  is an  $(\varepsilon, \delta)$ -domain and then apply Chua's extension theorem. Proving the  $(\varepsilon, \delta)$  property does not appear to be straightforward—one obstruction is that  $\Omega$  does not lie on one side of its boundary, as discussed in the Introduction.

## 5 Proof of the integration by parts formula

In this section we give a proof of the integration by parts formula, Theorem 2, which we now restate for the reader's convenience:

*Assume that  $\delta > 0$ , and let  $f \in C_c^1(E)$  and  $G \in C^1(E; \mathbf{M}^d)$ . If  $\delta \leq 1$ , assume that  $G(x)$  is tangent to  $\partial E$  at  $x$  for all  $x \in \partial E$ . If  $\delta < 1$ , assume in addition that  $G(x) \bullet x^{-\top}$  is locally bounded. Then*

$$\langle \nabla f, G \rangle = \langle f, -\nabla^* G \rangle. \quad (13)$$

Throughout the proof, let  $K$  be the (compact) support of  $f$ . For  $\varepsilon \geq 0$ , define

$$U_\varepsilon = \{x \in \mathbf{M}^d : \det x > \varepsilon\} \quad \text{and} \quad \nu(x) = -\frac{\nabla \det(x)}{\|\nabla \det(x)\|}, \quad x \in U_0.$$

For  $\varepsilon > 0$ ,  $U_\varepsilon$  has smooth boundary with outward unit normal  $\nu(x)$  at  $x \in \partial U_\varepsilon$ . For any smooth function  $h : U_0 \rightarrow \mathbf{R}$  such that the integrals are well-defined, the standard integration by parts formula yields, for each  $\varepsilon > 0$ ,

$$\begin{aligned} \int_{U_\varepsilon} \nabla f(x) \bullet G(x) h(x) dx &= \int_{\partial U_\varepsilon} f(x) h(x) G(x) \bullet \nu(x) d\sigma_\varepsilon(x) \\ &\quad - \int_{U_\varepsilon} f(x) \nabla \bullet (Gh)(x) dx, \end{aligned} \quad (14)$$

where  $\sigma_\varepsilon$  denotes the surface area measure on  $\partial U_\varepsilon$ .

*Case 1:*  $\delta > 1$ . Take  $h(x) = \det(x)^{\delta-1}$ . As  $\varepsilon \downarrow 0$ , the left side of (14) tends to  $\int_E \nabla f(x) \bullet G(x)m(dx)$  by dominated convergence. Let  $C > 0$  be such that  $|f(x)|\|G(x)\| \leq C$  for all  $x \in K$ . The absolute value of the boundary term is then dominated by

$$C\varepsilon^{\delta-1} \int_{\partial U_\varepsilon \cap K} d\sigma_\varepsilon(x),$$

using also that  $h(x) = \varepsilon^{\delta-1}$  for  $x \in \partial U_\varepsilon$ . It is easy to see that  $\sigma_\varepsilon(K)$  remains bounded as  $\varepsilon \downarrow 0$ , so we conclude that the boundary term vanishes in the limit. Consider now the second term on the right side of (14). The product rule yields

$$\nabla \bullet (Gh) = h\nabla \bullet G + G \bullet \nabla h.$$

By dominated convergence,  $\int_{U_\varepsilon} f(x)\nabla G(x)h(x)dx \rightarrow \int_E f(x)\nabla G(x)m(dx)$ . Moreover,  $G(x) \bullet \nabla h(x) = (\delta-1)G(x) \bullet \nabla \det(x) \det(x)^{\delta-2}$ , and since  $\det(x)^{\delta-2}dx$  is a Radon measure due to Theorem 1 and the fact that  $\delta > 1$ , we may again use dominated convergence to get

$$\int_{U_\varepsilon} f(x)G(x) \bullet \nabla h(x)dx \rightarrow (\delta-1) \int_E f(x)G(x) \bullet x^{-\top} m(dx).$$

(Here we used the equality  $\nabla \det(x) \det(x)^{\delta-2}dx = x^{-\top} m(dx)$ .) Assembling the pieces gives the desired formula (13).

*Case 2:*  $\delta = 1$ . We again take  $h(x) = \det(x)^{\delta-1} \equiv 1$ . Except for the boundary term, everything works as in the case  $\delta > 1$ , if we just note that  $\nabla h = 0$ . Letting  $C$  be a bound on  $|f(x)|$  over  $K$ , the boundary term is dominated by

$$C\sigma_\varepsilon(K) \sup_{x \in U_\varepsilon \cap K} G(x) \bullet \nu(x).$$

Using that  $G(x)$  is tangent to  $\partial E$  at every  $x \in \partial E$  it is not hard to show that the supremum tends to zero. Hence (13) is established.

*Case 3:*  $\delta < 1$ . Things are now a bit more complicated due to the fact that  $\det(x)^{\delta-1}$  blows up at  $\partial E$ . To get around this, for each  $n$  let  $\tau_n$  be a smooth, nondecreasing function satisfying the following properties:

$$\tau_n(t) \leq t \wedge n, \quad \tau_n(t) = t \text{ for } t \leq n-1, \quad \tau_n(t) = n \text{ for } t \geq n+1$$

$$0 \leq \tau'_n \leq 1, \quad \tau_n(t) \uparrow t \text{ as } n \rightarrow \infty.$$

In (14) we now take  $h = h_n$ , where  $h_n = \tau_n \circ w$  and  $w(x) = \det(x)^{\delta-1}$ . We first hold  $n$  fixed and let  $\varepsilon \downarrow 0$ . The left side of (14) converges to  $\int_E \nabla f(x) \bullet G(x)h_n(x)dx$  by dominated convergence. The boundary term on the right side will vanish by the same argument as in the case  $\delta = 1$ . The integrand in the second term on the right side is in fact bounded, since by the properties of  $\tau_n$ ,

$$\nabla h_n(x) = (\delta-1)(\tau'_n \circ w(x))\nabla \det(x) \det(x)^{\delta-2} \leq (\delta-1)\nabla \det(x)(n+1)^{\frac{\delta-2}{\delta-1}}.$$

Dominated convergence gives the limit  $\int_E f(x) \nabla \bullet (Gh_n)(x) dx$ . Combining these results and applying the product rule gives the formula

$$\int_E \nabla f(x) \bullet G(x) h_n(x) dx = - \int_E f(x) \left( \nabla \bullet G(x) h_n(x) + G(x) \bullet \nabla h_n(x) \right) dx.$$

The final step is to send  $n$  to infinity. The left side converges to  $\int_{\mathbb{E}} \nabla f(x) \bullet G(x) m(dx)$  by dominated convergence, since  $h_n = \tau_n \circ h \upharpoonright h$ . Similarly,  $\int_{\mathbb{E}} f(x) \nabla \bullet G(x) h_n(x) (dx) \rightarrow \int_{\mathbb{E}} f(x) \nabla \bullet G(x) m(dx)$ . Finally,

$$\begin{aligned} \int_K |G(x) \bullet \nabla h_n(x)| dx &= (\delta - 1) \int_K \tau'_n \circ w(x) \left| G(x) \bullet \frac{\nabla \det(x)}{\det(x)} \right| m(dx) \\ &\leq (\delta - 1) \int_K |G(x) \bullet x^{-\top}| m(dx). \end{aligned}$$

The right side is finite since  $G \bullet x^{-\top}$  is bounded on  $K$  by hypothesis, so dominated convergence yields  $\int_{\mathbb{E}} f(x) G(x) \bullet \nabla h_n(x) (dx) \rightarrow \int_{\mathbb{E}} f(x) G(x) \bullet \nabla h(x) (dx)$ , and hence the result.

## 6 Proof of the Muckenhoupt $A_1$ property

This section is devoted to the proof of Part (i) of Theorem 9. That is, we fix  $\alpha \in (-1, 0]$  and define

$$w(x) = |\det x|^\alpha,$$

where we set  $w(x) = +\infty$  if  $\det x = 0$ . We will prove the existence of a constant  $C > 0$ , depending only on  $d$  and  $\alpha$ , such that the inequality

$$\frac{1}{|B|} \int_B w(x) dx \leq C \inf_{x \in B} w(x) \tag{15}$$

holds for every ball  $B \subset \mathbf{M}^d$ . Define the set  $\mathbf{D}_+^d$  of diagonal matrices with nonnegative and ordered diagonal elements,

$$\mathbf{D}_+^d = \{\text{Diag}(\sigma) : \sigma \in \mathbf{R}^d, \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_d \geq 0\}.$$

The open ball centered at  $x \in \mathbf{M}^d$  with radius  $r > 0$  is denoted by  $B(x, r)$ . Its intersection with the nonsingular matrices is denoted by  $B_*(x, r)$ . That is,

$$B(x, r) = \{y \in \mathbf{M}^d : \|x - y\| < r\}, \quad B_*(x, r) = \{y \in B(x, r) : \det y \neq 0\}.$$

The proof of the Muckenhoupt property is quite long and involved (but nonetheless mostly elementary), due to the relatively complicated geometrical structure of the set  $\partial E$ , which is where the weight function becomes singular. The main idea is to change variables and integrate over the product space  $O(d) \times T(d)$  instead of  $\mathbf{M}^d$ . Unfortunately, balls in  $\mathbf{M}^d$  do not always map to balls (or comparable shapes) in  $O(d) \times T(d)$ , and this is where the

main complications arise. The resolution to this issue resides in Lemma 8 below, which relies on a detailed analysis of the mapping taking  $x$  to its  $QR$ -decomposition.

We start with a lemma that establishes an inequality similar to (15), where the balls  $B$  are replaced by sets of the form  $U \cdot K$ , with  $U \subset O(d)$  measurable and  $K \subset T(d)$  a cube.

**Lemma 7** *Let  $-1 < \alpha \leq 0$ . Then there is a constant  $C_1 > 0$ , depending only on  $d$  and  $\alpha$ , such that the inequality*

$$\int_{U \cdot K} w(x) dx \leq C_1 |U \cdot K| \inf_{x \in U \cdot K} w(x)$$

*holds for any measurable subset  $U \subset O(d)$  and any cube  $K \subset T(d)$ .*

**Proof.** Pick a cube  $K = \{R \in T(d) : R_{ij} \in I_{ij}, i \leq j\}$ , where the  $I_{ij}$  are intervals, and let  $U \subset O(d)$  be measurable. By Lemma 1 we have

$$\int_{U \cdot K} w(x) dx = \mu(U) \int_K \prod_{i=1}^d R_{ii}^{d-i+\alpha} dR = \mu(U) \prod_{i < j} |I_{ij}| \prod_{i=1}^d \int_{I_{ii}} t^{d-i+\alpha} dt,$$

and similarly  $|U \cdot K| = \mu(U) \prod_{i < j} |I_{ij}| \prod_{i=1}^d \int_{I_{ii}} t^{d-i} dt$ . We also have

$$\inf_{x \in U \cdot K} w(x) = \inf_{R \in K} \prod_{i=1}^d R_{ii}^\alpha = \prod_{i=1}^d \inf_{t \in I_{ii}} t^\alpha.$$

Hence the result follows from the following Claim:

*Let  $\alpha \in (-1, 0]$  and  $\beta \geq 0$ . Then there is a constant  $C_{\alpha, \beta}$  such that for every interval  $I \subset (0, \infty)$ , we have*

$$\int_I t^{\alpha+\beta} dt \leq C_{\alpha, \beta} \int_I t^\beta dt \inf_{t \in I} t^\alpha.$$

To prove the Claim it suffices to consider  $I = (a, b)$  with  $b > a \geq 0$ . We obtain:

$$\begin{aligned} \int_I t^{\alpha+\beta} dt &= \frac{1}{\alpha + \beta + 1} \left( b^{\beta+1} - \left( \frac{a}{b} \right)^\alpha a^{\beta+1} \right) b^\alpha \\ &\leq \frac{1}{\alpha + \beta + 1} \left( b^{\beta+1} - a^{\beta+1} \right) b^\alpha \\ &= \frac{\beta + 1}{\alpha + \beta + 1} \int_I t^\beta dt \inf_{t \in I} t^\alpha, \end{aligned}$$

as required. ■

Consider now balls  $B(\Sigma, r)$ , where the diagonal elements of  $\Sigma \in \mathbf{D}_+^d$  are either “large” (comparable to the radius  $r$ ) or zero. The following result reduces the proof that (15) holds for balls of this form to an application of Lemma 7. In the statement of condition (16) below, we use the convention that  $\sigma_0 = \infty$  and that  $i$  runs over  $\{0, \dots, d\}$ .

**Lemma 8** Suppose  $\Sigma \in \mathbf{D}_+^d$  and  $r > 0$  satisfy the following property, where  $\sigma \in \mathbf{R}^d$  is the vector of diagonal elements of  $\Sigma$ :

$$\begin{aligned} &\text{There is an index } n \in \{0, 1, \dots, d\} \text{ such that} \\ &\sigma_i > (5 + 18d)r \text{ for all } i \leq n, \text{ and } \sigma_i = 0 \text{ for all } i > n. \end{aligned} \quad (16)$$

Then there is a measurable subset  $U \subset O(d)$  and a cube  $K \subset T(d)$  such that

$$B_*(\Sigma, r) \subset U \cdot K \subset B_*(\Sigma, C_2 r),$$

where  $C_2$  is a positive constant that only depends on  $d$ .

**Proof.** The problem of finding the advertised constant  $C_2$  can be reduced to proving the following Claim, where  $e_1, \dots, e_d$  denote the canonical unit vectors in  $\mathbf{R}^d$ :

*There is a constant  $C_3$ , depending only on  $d$ , such that the following holds: For any  $x \in B_*(\Sigma, r)$ , let  $x = QR$  be its  $QR$ -decomposition, and let  $q_1, \dots, q_d$  be the columns of  $Q$ . Then the inequalities  $\|R - \Sigma\| < C_3 r$  and  $|q_i - e_i| < r\sigma_i^{-1}C_3$  hold for all  $i \in \{1, \dots, n\}$ , where  $n$  is the index from condition (16).*

Let us show how the statement of the lemma follows from this claim. Define  $K$  to be the cube in  $T(d)$  centered at  $\Sigma$  with side  $2C_3 r$ , i.e.

$$K = \{R \in T(d) : \max_{ij} |R_{ij} - \Sigma_{ij}| < C_3 r\},$$

and let  $U \subset O(d)$  be given by

$$U = \left\{ Q = (q_1, \dots, q_d) \in O(d) : |q_i - e_i| < \frac{r}{\sigma_i} C_3, i = 1, \dots, n \right\}.$$

The Claim then directly implies  $B_*(\Sigma, r) \subset U \cdot K$ . Next, for any  $x = QR \in U \cdot K$  the triangle inequality and the definition of  $K$  yield

$$\|x - \Sigma\| \leq \|R - \Sigma\| + \|(Q - I)\Sigma\| < \sqrt{d(d+1)/2} C_2 r + \|(Q - I)\Sigma\|.$$

Furthermore, since  $\sigma_i = 0$  for  $i > n$ , we have  $\|(Q - I)\Sigma\|^2 = \sigma_1^2 |q_1 - e_1|^2 + \dots + \sigma_n^2 |q_n - e_n|^2$ . We then deduce from the Claim that  $\|(Q - I)\Sigma\| < \sqrt{n} C_3 r$ , and consequently

$$\|x - \Sigma\| < C_2 r, \quad \text{where} \quad C_2 = \left( \sqrt{d(d+1)/2} + \sqrt{d} \right) C_3.$$

We are thus left with proving the Claim. Since it is vacuously true for  $n = 0$ , we can assume  $n \geq 1$ . The proof relies on a rather careful analysis of the Gram-Schmidt orthogonalization procedure for obtaining the  $QR$ -decomposition of a generic matrix  $x \in B_*(\Sigma, r)$ , so we briefly recall this procedure. To improve readability, we temporarily (for this proof only) adopt the notation  $\langle y, z \rangle = y^\top z$  for  $y, z \in \mathbf{R}^d$ . Fix  $x \in B_*(\Sigma, r)$  and let  $x_1, \dots, x_d$  be the columns of  $x$ . To obtain the  $QR$ -decomposition of  $x$ , one defines

$$u_1 = x_1, \quad q_1 = \frac{u_1}{|u_1|},$$



and, if  $q_1, \dots, q_{j-1}$  have been defined,

$$u_j = x_j - \sum_{i=1}^{j-1} \langle q_i, x_j \rangle q_i, \quad q_j = \frac{u_j}{|u_j|}. \quad (17)$$

The vectors  $q_1, \dots, q_d$  obtained in this way are the columns of  $Q$ , and  $R$  is given by  $R_{ij} = \langle q_i, x_j \rangle$ ,  $i \leq j$ .

We now proceed with the proof of the Claim. Recall that  $e_1, \dots, e_d$  are the canonical unit vectors in  $\mathbf{R}^d$ . Since  $x \in B(\Sigma, r)$ , we have  $x_i = \sigma_i e_i + h_i$ , where  $h_i$  is a vector in  $\mathbf{R}^d$  with  $|h_i| < r$ . Also let  $a = 5 + 18d$  denote the constant appearing in condition (16).

Fix  $j \in \{1, \dots, n\}$ , and suppose we have proved the following:

$$\text{For all } i \leq j-1 \text{ and all } k > i, \quad |R_{ik}| < 3r. \quad (18)$$

Then (17) and the inequalities  $|x_j| \geq \sigma_j - |h_j| \geq \sigma_j - r$  imply

$$|u_j| \geq \sigma_j - r - \sum_{i=1}^{j-1} |R_{ij}| \geq \sigma_j - r(1 + 3(j-1)). \quad (19)$$

Moreover, for  $k > j$  we use (18) and the fact that  $|\langle x_j, x_k \rangle| \leq \sigma_j r + \sigma_k r + r^2$  to get

$$|\langle u_j, x_k \rangle| \leq |\langle x_j, x_k \rangle| + \sum_{i=1}^{j-1} |R_{ij}| |R_{ik}| \leq \sigma_j r + \sigma_k r + r^2(1 + 9(j-1)).$$

It follows that

$$\begin{aligned} |R_{jk}| &= \frac{|\langle u_j, x_k \rangle|}{|u_j|} \leq r \frac{\sigma_j + \sigma_k + r(1 + 9(j-1))}{\sigma_j - r(1 + 3(j-1))} \\ &= r \frac{1 + \sigma_k/\sigma_j + (r/\sigma_j)(1 + 9(j-1))}{1 - (r/\sigma_j)(1 + 3(j-1))} \\ &\leq r \frac{2 + a^{-1}(1 + 9(j-1))}{1 - a^{-1}(1 + 3(j-1))}, \end{aligned}$$

where in the last step we used that  $\sigma_k \leq \sigma_j$  (since  $k > j$ ) and  $\sigma_j > ar$  (since  $j \leq n$ ). Since  $a = 5 + 18d$ , the right side is smaller than  $3r$ , as one readily verifies. We deduce that (18) holds with  $j$  replaced by  $j+1$ , and since it is vacuously true for  $j=1$  it follows by induction that it holds for all  $j \in \{1, \dots, n+1\}$ .

We now use this result to bound  $|R_{jj} - \sigma_j|$  for  $j \in \{1, \dots, n\}$ . To this end, write

$$|R_{jj} - \sigma_j| = \frac{1}{|u_j|} \left| |x_j|^2 - \sum_{i=1}^{j-1} \langle q_i, x_j \rangle^2 - \sigma_j |u_j| \right| \leq \frac{||x_j|^2 - \sigma_j |u_j|| + 9(j-1)r^2}{|u_j|},$$

using that  $\langle q_i, x_j \rangle^2 = |R_{ij}|^2 < 9r^2$  due to (18). Moreover, we have

$$||x_j|^2 - \sigma_j |u_j|| = |\sigma_j^2 - \sigma_j |u_j| + |h_j|^2 + 2\sigma_j \langle e_j, h_j \rangle| \leq \sigma_j |\sigma_j - |u_j|| + r^2 + 2\sigma_j r,$$

and by the reverse triangle inequality,

$$|\sigma_j - |u_j|| \leq |\sigma_j e_j - u_j| \leq r + 3(j-1)r. \quad (20)$$

Assembling the pieces and using the bound (19) gives

$$|R_{jj} - \sigma_j| \leq \frac{\sigma_j r(1 + 3(j-1)) + r^2 + 2\sigma_j r + 9(j-1)r^2}{\sigma_j - r(1 + 3(j-1))}.$$

Dividing the numerator and denominator by  $\sigma_j$  and using that  $\sigma_j > ar$ , we finally arrive at

$$\begin{aligned} |R_{jj} - \sigma_j| &\leq r \frac{1 + 3(j-1) + a^{-1} + 2 + a^{-1}9(j-1)}{1 - a^{-1}(1 + 3(j-1))} \\ &= r \frac{3 + 3(j-1) + a^{-1}(1 + 9(j-1))}{1 - a^{-1}(1 + 3(j-1))} \\ &< r \frac{14 + 12d}{3}. \end{aligned}$$

The only elements of  $R$  that remain to analyze are  $R_{ij}$  for  $i \leq j$  and  $j > n$ . But  $x_j = h_j$  for these  $j$ , so  $|R_{ij}| = |\langle q_i, x_j \rangle| \leq |h_j| < r$ . We are now able to estimate  $\|R - \Sigma\|$  as follows:

$$\|R - \Sigma\| \leq \sum_{i \leq j} |R_{ij}| \leq r \left( \frac{14 + 12d}{3} \times d + 3 \times \frac{n(n-1)}{2} + d(d-n) \right).$$

A bound solely in terms of  $d$  is then easily obtained. For instance, we may take

$$\|R - \Sigma\| \leq rd \left( \frac{14 + 12d}{3} + 3 \times \frac{(d-1)}{2} + d \right). \quad (21)$$

Let us now focus on bounding  $|q_j - e_j|$ ,  $j \in \{1, \dots, n\}$ . The calculations are similar to the ones used to bound  $|R_{jj} - \sigma_j|$  above, but slightly simpler. We have

$$\begin{aligned} |q_j - e_j| &= \frac{1}{|u_j|} \left| x_j - \sum_{i=1}^{j-1} \langle q_i, x_j \rangle q_i - |u_j| e_j \right| \\ &\leq \frac{1}{|u_j|} (|\sigma_j - |u_j|| + r + 3(j-1)r) \\ &\leq \frac{2}{|u_j|} (r + 3(j-1)r), \end{aligned}$$

using (20) in the last step. Using again (19) together with  $\sigma_j > ar$ ,

$$|q_j - e_j| \leq \frac{r}{\sigma_j} \times \frac{2 + 6(j-1)}{1 - a^{-1}(1 + 3(j-1))} < \frac{r}{\sigma_j} \times \frac{5 + 15d}{2}.$$

The Claim, and hence the lemma, is now proved, if for  $C_3$  we take the maximum of  $\frac{1}{2}(5 + 15d)$  and the constant in (21). ■

The next result implies that the proof of (15) for any ball with diagonal center reduces to an application of Lemma 8.

**Lemma 9** *Pick any  $\Sigma = \text{Diag}(\sigma) \in \mathbf{D}_+^d$  and  $r > 0$ . There is a matrix  $\Sigma' = \text{Diag}(\sigma') \in \mathbf{D}_+^d$  and a real number  $r' > 0$  that satisfy the condition (16) (with  $\sigma$  replaced by  $\sigma'$ , and  $r$  by  $r'$ ), such that*

$$B(\Sigma, r) \subset B(\Sigma', r') \quad \text{and} \quad r' \leq (6 + 18d)^d r.$$

**Proof.** Let  $a > 0$  be a constant to be determined later. Suppose for some index  $i \in \{0, \dots, d-1\}$ , we have  $\sigma_{d-i} > a(1+a)^i r$ . Let  $k$  be the smallest such index, and define

$$\sigma' = (\sigma_1, \dots, \sigma_{d-k}, 0, \dots, 0), \quad r' = r(1+a)^k.$$

Then, since  $\sigma_{d-i} \leq a(1+a)^i r$  for all  $i < k$ ,

$$\begin{aligned} |\sigma - \sigma'| &\leq \sigma_{d-k+1} + \dots + \sigma_d \\ &\leq r \left( a(1+a)^{k-1} + \dots + a(1+a) + a \right) \\ &= r(1+a)^k - r \\ &= r' - r. \end{aligned}$$

By the triangle inequality,  $B(\Sigma, r) \subset B(\Sigma', r')$ , where  $\Sigma' = \text{Diag}(\sigma')$ . Setting  $a = 5 + 18d$ , we see that  $\Sigma', r'$  satisfy condition (16) with  $n = d - k$ .

It remains to consider the case where  $\sigma_{d-i} \leq a(1+a)^i r$  for all  $i \in \{0, \dots, d-1\}$ . Then any  $x \in B(\Sigma, r)$  satisfies

$$\|x\| \leq r + r \left( a(1+a)^{d-1} + \dots + a(1+a) + a \right) = r(1+a)^d,$$

so  $B(\Sigma, r) \subset B(0, r(1+a)^d)$ . With  $r' = r(1+a)^d$  and  $n = 0$ , condition (16) is again satisfied for  $a = 5 + 18d$ . This finishes the proof. ■

The proof of (15) is now straightforward. Indeed, pick any ball  $B = B(x, r)$ , and let  $x = U\Sigma V^\top$  be a singular value decomposition of  $x$ . Then  $B(x, r) = U \cdot B(\Sigma, r) \cdot V^\top$ , and together with the invariance of Lebesgue measure under orthogonal transformations and the fact that  $\det y = \det(U^\top y V)$  for any  $y \in \mathbf{M}^d$ , this leads to the equalities

$$\int_{B(x, r)} w(x) dx = \int_{B(\Sigma, r)} w(x) dx,$$

$$|B(x, r)| = |B(\Sigma, r)|,$$

$$\inf_{x \in B(x, r)} w(x) = \inf_{x \in B(\Sigma, r)} w(x).$$

Consequently (and using that  $\partial E$  is a nullset), it suffices to prove (15) for  $B$  replaced by  $B_* = B_*(\Sigma, r)$  with  $\Sigma \in \mathbf{D}_+^d$ . We then have the following chain of inequalities, where we set  $B'_* = B_*(\Sigma', r')$  with  $\Sigma'$  and  $r'$  from Lemma 9, and where  $U$ ,  $K$ , and  $C_2$  are obtained

by applying Lemma 8 to  $\Sigma'$ ,  $r'$ .

$$\begin{aligned}
\int_{B_*} w(x) dx &\leq \int_{U \cdot K} w(x) dx && (B_* \subset B'_* \subset U \cdot K) \\
&\leq C_1 |U \cdot K| \inf_{x \in U \cdot K} w(x) && (\text{Lemma 7}) \\
&\leq C_1 |U \cdot K| \inf_{x \in B_*} w(x) && (B_* \subset U \cdot K) \\
&\leq C_1 |B_*(\Sigma', C_2 r')| \inf_{x \in B_*} w(x) && (U \cdot K \subset B_*(\Sigma', C_2 r')) \\
&= C_1 C_2^{d^2} (6 + 18d)^{d^3} |B_*| \inf_{x \in B_*} w(x). && (r' \leq (6 + 18d)^d r)
\end{aligned}$$

This proves that (15) holds with  $C = C_1 C_2^{d^2} (6 + 18d)^{d^3}$ .

## References

- [1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140. Academic Press, 2003.
- [2] J. Bertoin. Excursions of a  $\text{BES}_0(d)$  and its drift term ( $0 < d < 1$ ). *Probability Theory and Related Fields*, 84:231–250, 1990.
- [3] R. Bhatia and T. Jain. Higher order derivatives and perturbation bounds for determinants. *Linear Algebra and its Applications*, 431:2102–2108, 2009.
- [4] M.-F. Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.
- [5] S.-K. Chua. Extension theorems on weighted Sobolev spaces. *Indiana Math. J.*, 41:1027–1076, 1992.
- [6] C. Cuchiero, D. Filipovic, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *Annals of Applied Probability*, 21(2):397–463, Apr. 2011.
- [7] C. Donati-Martin, Y. Doumerc, H. Matsumoto, and M. Yor. Some properties of the wishart processes and a matrix extension of the hartman-watson laws. *Publications of the Research Institute For Mathematical Sciences*, 40(4):1385–1412, Dec. 2004.
- [8] D. Drusvyatsky and M. Larsson. Approximating functions on stratifiable sets. *Submitted*, 2012.
- [9] A. Eberle. *Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators*, volume 1718 of *Lecture notes in mathematics*. Springer, Berlin, 1999.
- [10] S. N. Ethier and T. G. Kurtz. *Markov Processes, Characterization and Convergence*. Wiley, 1986.
- [11] J. Faraut. *Analysis on Lie groups: an introduction*, volume 110. Cambridge University Press, Cambridge, UK, 2008.

- [12] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter studies in mathematics*. De Gruyter, Berlin, 2nd rev. and ext. ed edition, 2011.
- [13] J. Heinonen, T. Kilpeläinen, and O. Martio. *Nonlinear potential theory of degenerate elliptic equations*. Dover Publications, Mineola, NY, dover ed edition, 2006.
- [14] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [15] P. W. Jones. Quasiconformal mappings and extendability of functions in Sobolev spaces. *Acta Mathematica*, 147(1):71–88, 1981.
- [16] T. Kilpeläinen. Weighted Sobolev spaces and capacity. *Annales Academiæ Scientiarum Fennicæ, Series A. I. Mathematica*, 19:95–113, 1994.
- [17] A. Kufner and B. Opic. How to define reasonably weighted sobolev spaces. *Commentationes Mathematicae Universitatis Carolinae*, 25(3):537–554, 1984.
- [18] J. Lee. *Introduction to Smooth Manifolds*, volume 218. Springer, New York, 2003.
- [19] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293. Springer, Berlin, 3rd ed edition, 1999.
- [20] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales*. Cambridge University Press, Cambridge, U.K., 2nd ed edition, 2000.
- [21] A. Torchinsky. *Real-variable methods in harmonic analysis*, volume v. 123 of *Pure and applied mathematics*. Academic Press, Orlando, 1986.